Approximated Flux Boundary Conditions for Raviart-Thomas Finite Elements on Domains with Curved Boundaries and Applications to First-Order System Least Squares

von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr. rer. nat.
genehmigte Dissertation

von

Dipl.-Math. Fleurianne Bertrand
geboren am 06.06.1987 in Caen (Frankreich)

2014

Referent: Prof. Dr. Gerhard Starke
Universität Duisburg-Essen
Fakultät für Mathematik
Thea-Leymann-Straße 9
45127 Essen

Korreferent: Prof. Dr. James Adler
Department of Mathematics
Tufts University
Bromfield-Pearson Building
503 Boston Avenue
Medford, MA 02155

Korreferent: Prof. Dr. Joachim Escher
Institut für Angewandte Mathematik
Gottfried Wilhelm Leibniz Universität Hannover
Welfengarten 1
30167 Hannover

Tag der Promotion: 15.07.2014

Acknowledgement:
This work was supported by the German Research Foundation (DFG) under grant STA402/10-1.

Abstract

Optimal order convergence of a first-order system least squares method using lowest-order Raviart-Thomas elements combined with linear conforming elements is shown for domains with curved boundaries. Parametric Raviart-Thomas elements are presented in order to retain the optimal order of convergence in the higher-order case in combination with the isoparametric scalar elements. In particular, an estimate for the normal flux of the Raviart-Thomas elements on interpolated boundaries is derived in both cases. This is illustrated numerically for the Poisson problem on the unit disk. As an application of the analysis derived for the Poisson problem, the effect of interpolated interface condition for a stationary two-phase flow problem is then studied.

Keywords: Raviart-Thomas, Curved Boundaries, First-Order System Least Squares

Zusammenfassung

Für Gebiete mit gekrümmten Rändern wird die optimale Konvergenzordnung einer Least Squares finite Elemente Methode für Systeme erster Ordnung mit Raviart-Thomas Elementen niedrigster Ordnung und linearen konformen Elementen gezeigt. Um die optimale Konvergenzordnung im Fall höherer Ordnung mit den isoparametrischen skalaren Elementen zu behalten, werden parametrische Raviart-Thomas Elemente eingeführt. Insbesondere wird eine Abschätzung für deren Flüsse in Richtung der Normalkomponente auf dem interpolierten Rand für beide Fälle hergeleitet. Am Beispiel des Poisson Problems auf dem Einheitskreis wird die Least Squares Methode für gekrümmte Ränder numerisch verdeutlicht. Als Anwendung der Analyse für das Poisson Problem dient die Untersuchung des Effekts der interpolierten Grenzkurven auf einem stationären Stokes Zwei-Phasen Problem.

Schlagwörter: Raviart-Thomas, gekrümmte Ränder, Least Squares finite Elemente Methode für Systeme erster Ordnung

Résumé

La convergence d´ordre optimal d´une méthode des éléments finis Least Squares pour un systeme de premier ordre est prouvée pour des domaines à bords courbes avec les éléments de Raviart-Thomas d´ordre le plus bas et les éléments finis conformes lineaires. Pour conserver un ordre de convergence optimal dans le cas d´éléments finis isoparamétriques d´ordre plus élevé, les éléments Raviart-Thomas paramétriques sont introduits. En particulier, leur flux dans la direction normale sur le bord interpolé est estimé dans les deux cas. L´exemple du problème de Poisson sur le disque unité illustre ce résultat. L´analyse dérivée pour le problème de Poisson est appliquée à un problème stationaire de flux à deux phases avec conditions interpolées sur l´interface.

Contents

Introduction

In this thesis, the first-order system least squares methods on domains with curved boundaries is analyzed. Thus, scalar and vector unknowns are approximated simultaneously but the inf-sup condition typically arising in mixed methods (see [11]) is not necessary ([34]). Further, an a posteriori error estimator is inherent in the least squares method and the linear equation system of the discrete problem is positive definite, but in return they have poor conservation properties. Comprehensive theory and computational aspects of least squares finite-element methods are given in [9], and [15] proves error bounds for second-order partial differential equations (see also [16] for the effect of additional constraint and boundary conditions). More details about the least squares mixed finite-elements in relation to standard and mixed finite-elements are given in [13] (see also [10] for the connection to the Dirichlet and Kelvin principles).

For scalar elements in the lowest-order case, it is well-known (see [12]) that a polygonal approximation of the boundary is sufficient to retain the optimal order of convergence. Further in the higher-order case, the isoparametric framework ensures a more accurate representation of the boundary by the same finite-element space and leads, therefore, to the optimal order of convergence as well (see [14] and [37]). A Matlab implementation of the quadratic order isoparametric finite-element method for the Laplace equation in two dimensions is given in [5]. However, for Raviart-Thomas elements where Neumann boundary conditions are imposed on the normal flux, this is more complicated as the normal flux has to be estimated on the interpolated boundary. Furthermore, the vector-valued finite-element space cannot be used to approximate the boundary and a parametric space is needed. These spaces and the corresponding Piola-mapping are introduced in [30] and [36]. Implementation of the lowest-order Raviart-Thomas elements are given in [4].

In order to derive a convergence analysis of Raviart-Thomas elements on curved domains in the context of first-order system least squares formulations, their normal flux on interpolated boundaries is estimated. Optimal order of convergence is shown in the lowest-order case if a polygonal approximation of the boundary is used. Parametric Raviart-Thomas elements are introduced, which retain the optimal order of convergence in the higher-order case.

This estimate for the normal flux on interpolated boundary is useful for two-phase problems, where the interface condition consists in a relation between the normal components of the two velocity fields, see for example water-mud interaction in [22]. A first-order system least squares formulation for equations of two-phase flow is given in [2]. A further example is the coupled (generalized Newtonian) Stokes-Darcy flow which is considered in [20], [33] and [32]).

For the purpose of exposition, the two-dimensional Poisson equation with homogenous boundary conditions is considered. A two-phase flow example introduces the way to control the

additional error caused by the inhomogeneous boundary conditions for the normal component of the approximated fields. [38] and [39] are related works for the treatment of the inhomogeneous boundary conditions. The first chapter of the present work is an introduction to the least squares finite-element method. The basis spaces of the theory of the weak formulation of partial differential equations, i.e. the Sobolev spaces are introduced and the necessary inequalities to prove the further results on these spaces are stated. Further, the used finite-element spaces are presented.

The second chapter describes the construction of a curved triangulation for curved domains and of the mapping between the polygonal domain and the curved domain. The finite-element spaces on the curved domain are introduced and an estimate for its normal flux on the interpolated boundary is derived.

The third chapter uses this estimate in order to bound the least squares functional from below and to show that the use of the polygonal approximation of the boundary derived in the second chapter is sufficient to retain the optimal order of convergence for the lowest-order elements but does not lead to an optimal order of convergence in the higher-order case.

The fourth chapter extends the study of the least squares method on domains with curved boundaries to the higher-order case. Therefore, the curved domain is approximated by a domain with piecewise polynomial boundary. The parametric Raviart-Thomas finite-elements on this approximated domain are introduced and interpolation operators for these finite-element spaces are presented. Further, the normal flux of the parametric Raviart-Thomas elements on the interpolated boundary is estimated.

This result is used in the fifth chapter for the error analysis of the first-order system least squares finite-element approximation. The main result states that the use of the polynomial approximation is sufficient to retain the optimal order of convergence for the parametric elements and is illustrated by a numerical example at the end of this chapter.

The sixth chapter applies the derived theory of the least squares finite-element on domains with curved boundary to a stationary two-phase flow model with interface condition on the normal flux and the seventh chapter states considerations for the three-dimensional case.

The closing chapter gives an outlook and a short conclusion.

Chapter 1

Least Squares Finite Element Methods for First-Order Systems

In this chapter, an introduction in the Least Squares Finite Element Methods (LSFEM) is given, where $\Omega \in \mathbb{R}^2$ is assumed to be a bounded polygonal domain with a piecewise linear boundary $\Gamma = \partial\Omega$. The first two sections present the idea of the least squares methodology to approximate the solution of the first-order system $\mathcal{Q}u = f$, where Q is assumed to be a linear first-order differential operator from a Hilbert space $\mathcal{V}$ into an L^2 product space.

The third section consists of an introduction of the Sobolev spaces that are fundamental in the study of partial differential equations. Note that the results stated in this section hold for a bounded Lipschitz domain, this means in particular for the polygonal domain assumed in this first chapter. The following sections take a closer look at the Sobolev spaces used in this work as a solution space.

Then, in Section 1.5, the previous theory is then applied on the Poisson problem and the subsections 1.5.1 and 1.5.2 present the finite-elements used for its approximation.

1.1 Least Squares Principles and Weak Formulation

The least squares finite-element method for a linear first-order system (see [9]) is based on minimizing the L^2 norm of the residuals in the first-order differential equations, that is the least squares functional $\mathcal{F}_{\mathcal{Q}}(v; f) = \|\mathcal{Q}v - f\|^2_{0,\Omega}$. The solution u is therefore defined as

$$u = \arg\min_{v \in \mathcal{V}} \mathcal{F}_{\mathcal{Q}}(v; f) \ . \tag{1.1}$$

If $(\cdot,\cdot)_{\mathcal{V}}$ denotes the inner product of the Hilbert space $\mathcal{V}$, the first-order necessary condition for the problem (1.1) is the following variational problem:

$$\text{Find } u \in \mathcal{V} \text{ such that } (\mathcal{Q}u, \mathcal{Q}v)_{0,\Omega} = (f, \mathcal{Q}v)_{0,\Omega} \ \forall v \in \mathcal{V} \ . \tag{1.2}$$

$\mathcal{V}$ has to be chosen such that a unique solution exists, in particular such that the correct boundary conditions are set. A sufficient condition therefore is the following energy balance for some positive constants α_1 and α_2:

$$\alpha_1 \|v\|_{\mathcal{V}} \leq \|\mathcal{Q}v\|_{0,\Omega} = \mathcal{F}_{\mathcal{Q}}(v; 0)^{\frac{1}{2}} \leq \alpha_2 \|v\|_{\mathcal{V}} \quad v \in \mathcal{V} \tag{1.3}$$

More theory about the energy balance is given in [9]. Here, just note that the energy balance implies that the Lax-Milgram theorem of the standard finite-element theory can be applied on

(1.2). Therefore, for any $f \in L^2(\Omega)$, the least squares minimization problem (1.1) has a unique solution $u \in \mathcal{V}$ such that

$$\mathcal{F}_{\mathcal{Q}}(u; f) = 0 \ . \tag{1.4}$$

In Section 1.5 the energy balance is derived for the Poisson problem. In order to simplify the notation there and in the next of this work, the following notations are introduced:

- $a(\xi) \lesssim b(\xi)$ and $b(\xi) \gtrsim a(\xi)$ means that there exists a positive constant C such that $a(\xi) \leq C b(\xi)$ holds for all admissible ξ.
- $a(\xi) \eqsim b(\xi)$ means that $a(\xi) \lesssim b(\xi)$ and $a(\xi) \gtrsim b(\xi)$ hold.

1.2 Least Squares Finite Element Method

In order to describe the least squares finite-element method, a finite dimensional subspace $\mathcal{V}_h$ of $\mathcal{V}$ with respect to a discretisation of Ω is needed. In this work, this means a triangulation $\mathcal{T}_h$ of Ω, where h denotes the maximal diameter of the triangles in the triangulation. Further, a shape regular family of triangulations $\{\mathcal{T}_h\}$ will be considered. It means, that there exists a number $\lambda > 0$ such that every element $T \in \mathcal{T}_h$ contains a circle of radius ρ_T with

$$\rho_T \geq \frac{h_T}{\lambda} \ , \tag{1.5}$$

where h_T is the diameter of T (see [12]). For an integer $k \geq 0$, $\Lambda_h : \mathcal{V} \to \mathcal{V}_h$ denotes then an h^k-interpolation operator, such that it holds:

$$\|v - \Lambda_h(v)\|_{\mathcal{V}} \lesssim h^k \|v\|_{\mathcal{V}} \ . \tag{1.6}$$

Note that usually for this estimate, additionally regularity assumptions have to hold. These will be given in Sections 1.5.1 and 1.5.2 for the different subspaces. The discrete solution u_h of the problem (1.1) is given by

$$u_h = \underset{v_h \in \mathcal{V}_h}{\arg\min} \, \mathcal{F}_{\mathcal{Q}}(v_h; f) \ . \tag{1.7}$$

Due to the subspace properties, (1.3) holds for all $v_h \in \mathcal{V}_h$ and the variational formulation

$$\text{Find } u_h \in \mathcal{V}_h \text{ such that } (\mathcal{Q}u_h, \mathcal{Q}v_h)_{0,\Omega} = (f, \mathcal{Q}v_h)_{0,\Omega} \qquad \forall v_h \in \mathcal{V}_h \ , \tag{1.8}$$

has a unique minimizer that solves (1.7). Combining the minimizing property of the least squares method (1.7), the ellipticity (1.3) of the least squares functional and the interpolation property leads to

$$\mathcal{F}_{\mathcal{Q}}(u_h)^{\frac{1}{2}} \leq \mathcal{F}_{\mathcal{Q}}(\Lambda_h(u))^{\frac{1}{2}} \lesssim \|u - \Lambda_h(u)\|_{\mathcal{V}} \lesssim h^k \|u\|_{\mathcal{V}} \ . \tag{1.9}$$

This means, that the LSFEM converges to the solution u at least as fast as the interpolation operator. The energy balance (1.3) and the linearity of $\mathcal{Q}$ directly imply that the error is a lower and an upper bound for the least squares functional:

$$\|u - u_h\|_{\mathcal{V}}^2 \lesssim \|\mathcal{Q}(u - u_h)\|^2 = \mathcal{F}_{\mathcal{Q}}(u_h) \lesssim \|u - u_h\|_{\mathcal{V}}^2 \ . \tag{1.10}$$

On one hand combining (1.9) and (1.10) implies the a priori error estimate

$$\|u - u_h\|_{\mathcal{V}} \lesssim h^k \|u\|_{\mathcal{V}} \ . \tag{1.11}$$

On the other hand (1.10) allows the usage of the least squares functional as an a posteriori error estimator, that is a major avantage of the LSFEM. Usually, $\mathcal{V}_h$ consists on piecewise polynomials of degree k subject to the triangulation $\mathcal{T}_h$:

$$\mathcal{V}_h \subset \left(\{v_h \in L^2(\Omega) : v_h|_T \in \mathcal{P}_k(T) \quad \forall T \in \mathcal{T}_h\}\right)^{\bar{d}}, \tag{1.12}$$

where $\mathbb{P}_k(T)$ denotes the space of polynomials of degree k on a triangle T for an integer $k \geq 0$ and $\bar{d}$ denotes the dimension of the variables that are approximated with $\mathcal{V}_h$. $\mathcal{V}_h$ as a conforming finite element space has to satisfy two requirements. First, it is locally defined for an element $T \in \mathcal{T}_h$ with the help of a so-called finite-element, and second, the finites elements have to be composed such that the globally defined finite-element space is a subspace of $\mathcal{V}$. For $T \in \mathcal{T}_h$, a finite-element is defined (see [18]) as the triple $(T, \mathcal{P}_T, \mathcal{N}_T)$, where $\mathcal{P}_T$ is a finite dimensional linear space of functions called the space of shape functions and $\mathcal{N}_T = \{l_{1,T}, l_{2,T}, ..., l_{N,T}\}$ is a basis for $\mathcal{P}_T'$, called the set of nodal variables. $\Sigma_T = \{\varphi_{1,T}, \varphi_{2,T}, ..., \varphi_{N,T}\}$ denotes the basis of $\mathcal{P}_T$ dual to $\mathcal{N}_T$. Its elements are called degrees of freedom. The local interpolant can then be given with

$$\Lambda_T v = \sum_{i=1}^{k} l_{i,T}(v)\varphi_{i,T} \tag{1.13}$$

and the global interpolant is then defined by

$$\Lambda_{\mathcal{T}_h} v|_T = \Lambda_T v \qquad \forall\ T \in \mathcal{T}_h \ . \tag{1.14}$$

The sections 1.5.1 and 1.5.2 present the finite-elements used in this work, that are the Lagrange element and the Raviart-Thomas element.

1.3 Basic Theory of Sobolev Spaces

In the previous sections, the least squares method was introduced for an abitrary Hilbert space. The aim of this section is to define the Hilbert spaces used in this work and more generally the Sobolev spaces, that are the basic spaces of the theory of the weak formulation of partial differential equations. For any positive integer m and $1 \leq p \leq \infty$ the Sobolev space W_p^m can be defined in two ways (see [1]):

$$\begin{aligned} W_p^m(\Omega) &= \text{the completion of } \{v \in C^m(\Omega) \ : \ \|v\|_{m,p,\Omega} < \infty\} \\ &\qquad \text{with respect to the norm } \|v\|_{m,p,\Omega} \ , \\ W_p^m(\Omega) &= \{v \in L^p(\Omega) \ : \ D^\alpha v \in L^p(\Omega) \text{ for } 0 \leq |\alpha| \leq m\} \ , \end{aligned} \tag{1.15}$$

where

$$\|v\|_{m,p,\Omega} = \begin{cases} \left(\sum\limits_{0 \leq |\alpha| \leq m} \|D^\alpha v\|_p^p\right)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \max\limits_{0 \leq |\alpha| \leq m} \|D^\alpha v\|_\infty & p = \infty \end{cases} \tag{1.16}$$

for admissible v. Let $|v|_{m,p,\Omega}$ denote the corresponding seminorm:

$$|v|_{m,p,\Omega} = \begin{cases} \left(\sum\limits_{|\alpha|=m} \|D^\alpha v\|_p^p \right)^{\frac{1}{p}} & \text{if } 1 \le p < \infty \\ \max\limits_{|\alpha|=m} \|D^\alpha v\|_\infty & p = \infty \end{cases} . \tag{1.17}$$

For $p = 2$, the norm $\|v\|_{m,2}$ is the norm induced by the inner product

$$(v_1, v_2)_{W_2^m(\Omega)} = (v_1, v_2)_{m,\Omega} = \sum_{0 \le |\alpha| \le m} (D^\alpha v_1, D^\alpha v_2)_{L^2(\Omega)} \tag{1.18}$$

such that $W^{m,2}(\Omega)$ is a Hilbert space that is usually denoted by $H^m(\Omega)$. In this case, the norm $\|v\|_{m,2,\Omega}$ is simply denoted by $\|v\|_{m,\Omega}$. Important inequalities for the next of this work are given in the following Sobolev imbedding theorem, where $X \to Y$ means that $X \subset Y$ (normed space equipped with the norm $\|\cdot\|_X$) is imbedded in the normed space Y (equipped with the norm $\|\cdot\|_Y$), i.e. the identity operator $\mathrm{id} : X \to Y$ is continuous. This continuity of the identity operator is equivalent to the existence of a constant C such that

$$\|\mathrm{id}(x)\|_Y \le C\|x\|_X \tag{1.19}$$

Theorem 1.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with a Lipschitz continuous and piecewise C^1-boundary Γ. Let $j \ge 0$ and $m \ge 1$ be integers.*

Due to $L^q(\Omega) \subset L^p(\Omega)$ it holds	$W_q^j(\Omega) \to W_p^j(\Omega)$	*for $1 \le p \le q \le \infty$.*	(1.20a)
If $mp > n$, then	$W_p^{j+m}(\Omega) \to W_q^j(\Omega)$	*for $1 \le p \le q \le \infty$*	(1.20b)
If $mp = n$, then	$W_p^{j+m}(\Omega) \to W_q^j(\Omega)$	*for $1 \le p \le q < \infty$*	(1.20c)
Moreover,	$W_1^{j+n}(\Omega) \to W_q^j(\Omega)$	*for $1 \le q \le \infty$*	(1.20d)
and	$W_p^{j+m}(\Omega) \to W_p^j(\Omega)$	*for $1 \le p \le \infty$*	(1.20e)

Proof: Combine theorem 4.12 for $k = 1$ in [1] with the fact that the Lipschitz assumption implies the uniform cone condition. (1.20b) is the Case A and (1.20c) is the Case B. (1.20d) is the Case A with $p = 1$ and (1.20e) is all cases together for the special case $p = q$. □

As mentioned in the previous section, it is important to be able to set boundary conditions on these Hilbert spaces. First, note the following trace imbedding theorem:

Theorem 1.2. *Let Ω be a bounded domain with a Lipschitz continuous and piecewise C^1-boundary Γ. There exists a constant C_T with*

$$\|v\|_{0,\Gamma} \le C_T \|v\|_{1,\Omega} \quad \forall\ v \in H^1(\Omega) . \tag{1.21}$$

Proof: Combine the theorems 5.36 and 5.28 in [1] with the fact that the Lipschitz assumption implies the uniform cone condition. □

The embedding from $H^1(\Omega)$ to $L^2(\Gamma)$ mapping u on $u_{|\Gamma}$ described by the inequality (1.21) is usually denoted with γ_0 and allows the following definition of the space of functions in $H^1(\Omega)$ that vanish at Γ in the sense of trace (see [31]):

$$H_0^1(\Omega) = \{u \in H^1(\Omega) \ :\ \gamma_0(u) = 0\} . \tag{1.22}$$

Note that the mapping γ_0 is not surjective. For a bounded domain Ω with Lipschitz boundary, the space of the traces $v_{|\Gamma}$ of functions $v \in H^m(\Omega)$ is denoted by $W_2^{m-\frac{1}{2}}(\Gamma) = H^{m-\frac{1}{2}}(\Gamma)$.

There are many ways to define the fractional order Sobolev space $H^{m-\frac{1}{2}}(\Gamma)$ for $m \geq 1$, and the definition chosen here is based on the so-called $1 - \frac{1}{2m}$ interpolation between the spaces $L^2(\Gamma)$ and $H^m(\Gamma)$:

$$\begin{aligned} H^{m-\frac{1}{2}}(\Gamma) &= \left(L^2(\Gamma), H^m(\Gamma)\right)_{1-\frac{1}{2m},2,J} \\ &= \{q \in L^2(\Gamma) \ : \ q = \int_0^\infty f(t)\frac{dt}{t} \text{ for some } f : \mathbb{R}^+ \to H^m(\Gamma) \text{ with } \int_0^\infty \|f(t)\|_{0,\Gamma}\frac{dt}{t} < \infty \\ &\qquad \text{and such that } \left(\int_0^\infty \left(t^{-s}\max\left(\|q\|_{0,\Gamma}, t\|q\|_{m,\Gamma}\right)\right)^{\frac{1}{2}}\frac{dt}{t}\right) < \infty\} \ . \end{aligned}$$

A consequence of this interpolation definition is the following interpolation theorem (theorem 7.23 in [1]):

Theorem 1.3. *Let $T : L^2(\Gamma) \to \mathbb{R}$ be a bounded linear operator and let C_0 and C_m be two positive constants such that it holds*

$$\begin{aligned} |Tq| &\leq C_0\|q\|_{0,\Gamma} && \forall q \in L^2(\Gamma) \\ \text{and } |Tq| &\leq C_m\|q\|_{m,\Gamma} && \forall q \in H^m(\Gamma). \end{aligned}$$

Then, it exists a positive constant $C_{m-\frac{1}{2}} \leq C_I C_m \left(\frac{C_0}{C_m}\right)^{\frac{1}{2m}}$ such that it holds

$$|Tq| \leq C_{m-\frac{1}{2}}\|q\|_{m-\frac{1}{2},\Gamma} \qquad \forall q \in H^{m-\frac{1}{2}}(\Gamma)$$

with $C_I \geq 1$ independent of T.

The embedding γ_0 from $H^1(\Omega)$ to $H^{\frac{1}{2}}$ is now surjective, such that the following theorem holds.

Theorem 1.4. *Let Ω be a bounded domain with a Lipschitz continuous and piecewise C^1-boundary Γ. There exists a constant C_T with*

$$\|v\|_{\frac{1}{2},\Gamma} \leq C_T\|v\|_{1,\Omega} \quad \forall \ v \in H^1(\Omega) \ . \tag{1.23}$$

A further important inequality on Sobolev spaces is the following Poincaré inequality:

Lemma 1.1. *Let Ω be an open bounded Lipschitz domain. Then there exists a constant C_Ω, depending only on Ω such that*

$$\|v\|_{1,\Omega} \leq C_\Omega\left(\|\nabla v\|_{0,\Omega} + |(v,1)_{0,\Omega}|\right) \quad \forall v \in H^1(\Omega) \ . \tag{1.24}$$

Proof: See [14] and [31].
Note that in particular for all $v \in \dot{H}^1(\Omega) = \{v \in H^1(\Omega) : (v,1)_{0,\Omega} = 0\}$ it holds

$$\|v\|_{0,\Omega}^2 \leq \|v\|_{1,\Omega}^2 \leq C_\Omega\|\nabla v\|_{0,\Omega}^2 \ . \tag{1.25}$$

Further, combining the trace inequality and the Poincaré inequality leads to

$$\|v\|_{\frac{1}{2},\Gamma} \leq C_\Omega C_T\|\nabla v\|_{1,\Omega} \quad \forall \ v \in \dot{H}^1(\Omega) \ . \tag{1.26}$$

1.4 The Space $H(\mathbf{div}, \Omega)$

This section states results concerning the space of vector functions having a square-integrable divergence:

$$H(\mathrm{div}, \Omega) = \{\mathbf{v} \in \left(L^2(\Omega)\right)^2 \ : \ \mathrm{div}\ \mathbf{v} \in L^2(\Omega)\}. \tag{1.27}$$

This is a Hilbert space for the norm

$$\|\mathbf{v}\|_{\mathrm{div},\Omega} = \left(\|\mathbf{v}\|^2_{0,\Omega} + \|\mathrm{div}\ \mathbf{v}\|^2_{0,\Omega}\right)^{\frac{1}{2}}. \tag{1.28}$$

Let $\mathbf{n}$ denote the outward pointing normal vector on Γ and $H^{-\frac{1}{2}}(\Gamma)$ the dual of the fractional order Sobolev space $H^{\frac{1}{2}}(\Gamma)$ defined in the last section. $H^{-\frac{1}{2}}(\Gamma)$ is equipped with the negative norm

$$\|v\|_{-\frac{1}{2},\Gamma} = \sup_{w \in H^{\frac{1}{2}}(\Gamma),\ w \neq 0} \frac{\langle v, w\rangle_{0,\Gamma}}{\|w\|_{\frac{1}{2},\Gamma}}. \tag{1.29}$$

A major property of functions in $H(\mathrm{div}, \Omega)$ is that they have well-defined normal components on Γ, as stated in the following theorem.

Theorem 1.5. *Let Ω be a bounded domain with a Lipschitz continuous boundary Γ. Then, it exists a linear continuous mapping γ_n from $H(div, \Omega)$ onto $H^{-\frac{1}{2}}(\Gamma)$ with*

$$\gamma_n(\mathbf{v}) = \mathbf{v}_{|\Gamma} \cdot \mathbf{n}. \tag{1.30}$$

In particular, it holds

$$\|\mathbf{v} \cdot \mathbf{n}\|_{-\frac{1}{2},\Gamma} \leq C_N \|\mathbf{v}\|_{\mathrm{div},\Omega}. \tag{1.31}$$

Further, the Green's formula

$$(\mathbf{v}, \nabla q)_{0,\Omega} + (\mathrm{div}\ \mathbf{v}, q)_{0,\Omega} = \langle q, \gamma_n(\mathbf{v})\rangle_\Gamma \tag{1.32}$$

holds for functions $\mathbf{v} \in H(div, \Omega)$ and $q \in H^1(\Omega)$.

Proof: See [31].

1.5 Least Squares Method for the Poisson Problem

The aim of this section is to apply the LSFEM to the problem considered in the next chapters, that is the Poisson problem:

$$\begin{aligned} \mathrm{div}\ \mathbf{u} &= f \\ \mathbf{u} + \nabla p &= \mathbf{0}\ , \end{aligned} \tag{1.33}$$

with given $f \in L^2(\Omega)$ subject to Neumann boundary conditions $\mathbf{n} \cdot \mathbf{u} = 0$ on $\Gamma = \partial\Omega$.
In order to get a unique solution the solution has to be normalized, that means the following additional condition has to hold:

$$\int_\Omega p = 0\ . \tag{1.34}$$

The associated least squares functional is then given by

$$\mathcal{F}(\mathbf{u},p) = \|\text{div } \mathbf{u} - f\|_{0,\Omega}^2 + \|\mathbf{u} + \nabla p\|_{0,\Omega}^2, \tag{1.35}$$

for $(\mathbf{u},p) \in H_\Gamma(\text{div},\Omega) \times \dot{H}^1(\Omega)$, with the space $H_\Gamma(\text{div},\Omega) = \{\mathbf{v} \in H(\text{div},\Omega) : \langle \mathbf{v}\cdot\mathbf{n}, q\rangle_{0,\Gamma} = 0 \ \forall q \in H^{\frac{1}{2}}(\Omega)\}$. The variational formulation (1.2) becomes

$$\begin{aligned}&\text{Find } (\mathbf{u},p) \in H_\Gamma(\text{div},\Omega) \times \dot{H}^1(\Omega) \text{ such that}\\ &(\text{div } \mathbf{u}, \text{div } \mathbf{v})_{0,\Omega} + (\mathbf{u} + \nabla p, \mathbf{v} + \nabla q)_{0,\Omega} = (f, \text{div } \mathbf{v})_{0,\Omega} \ \forall \ (\mathbf{v},q) \in H_\Gamma(\text{div},\Omega) \times \dot{H}^1(\Omega)\ ,\end{aligned} \tag{1.36}$$

and the norm equivalency that has to hold for all $(\mathbf{v},q) \in H_\Gamma(\text{div},\Omega) \times \dot{H}^1(\Omega)$ is:

$$\|\mathbf{v}\|_{H_\Gamma(\text{div},\Omega)}^2 + \|q\|_{\dot{H}^1(\Omega)}^2 \lesssim \|\text{div } \mathbf{v}\|_{0,\Omega)}^2 + \|\mathbf{v} + \nabla q\|_{0,\Omega}^2 \lesssim \|\mathbf{v}\|_{H_\Gamma(\text{div},\Omega)}^2 + \|q\|_{\dot{H}^1(\Omega}^2\ . \tag{1.37}$$

The upper bound follows directly from the definitions of the norms:

$$\|\mathbf{v}\|_{H_\Gamma(\text{div},\Omega)}^2 + \|q\|_{\dot{H}^1(\Omega)}^2 = \|\mathbf{v}\|_{\text{div},\Omega}^2 + \|q\|_{1,\Omega}^2, \tag{1.38}$$

while the idea to get the lower bound is to expand the term $\|\mathbf{v} + \nabla q\|_{0,\Omega}^2$ and to distribute the mixed term on the other terms and in particular on the $\|\text{div } \mathbf{v}\|_{0,\Omega}^2$ term from the least squares functional. A useful inequality therefore is

$$\pm ab \leq \frac{a^2}{2C} + \frac{Cb^2}{2}\ , \tag{1.39}$$

that holds for any constant $C > 0$. Further, note that integration by parts with the Neumann boundary conditions leads to

$$(\mathbf{v}, \nabla q)_{0,\Omega} = -(\text{div } \mathbf{v}, q)_{0,\Omega} \tag{1.40}$$

and therefore

$$\begin{aligned}-|2(\mathbf{v},\nabla q)_{0,\Omega}| &= -|-2\alpha(\text{div } \mathbf{v}, q)_{0,\Omega} + 2(1-\alpha)(\mathbf{v},\nabla q)_{0,\Omega}|\\ &\geq -\frac{1}{2}\|\text{div } \mathbf{v}\|_{0,\Omega}^2 - 2\alpha^2\|q\|_{0,\Omega}^2 - (1-\alpha)\|\mathbf{v}\|_{0,\Omega}^2 - (1-\alpha)\|\nabla q\|_{0,\Omega}^2\ ,\end{aligned}$$

for $\alpha \in [0,1]$ and with the Poincaré inequality (1.25):

$$-|2(\mathbf{v},\nabla q)_{0,\Omega}| \gtrsim -\frac{1}{2}\|\text{div } \mathbf{v}\|_{0,\Omega}^2 - (1-\alpha+2\alpha^2 C_\Omega)\|\nabla q\|_{0,\Omega}^2 - (1-\alpha)\|\mathbf{v}\|_{0,\Omega}^2. \tag{1.41}$$

Then it holds:

$$\begin{aligned}\|\text{div } \mathbf{v}\|_{0,\Omega}^2 + \|\mathbf{v} + \nabla q\|_{0,\Omega}^2 =& \|\text{div } \mathbf{v}\|_{0,\Omega}^2 + \|\mathbf{v}\|_{0,\Omega}^2 + \|\nabla q\|_{0,\Omega}^2 - 2(\mathbf{v},\nabla q)_{0,\Omega}\\ \gtrsim& \frac{1}{2}\|\text{div } \mathbf{v}\|_{0,\Omega}^2 + \alpha(1-2\alpha C_\Omega)\|\nabla q\|_{0,\Omega}^2 + \alpha\|\mathbf{v}\|_{0,\Omega}^2.\end{aligned}$$

Choosing $\alpha = \min\left(1, \frac{1}{2C_\Omega}\right)$ leads to

$$\begin{aligned}\|\text{div } \mathbf{v}\|_{0,\Omega}^2 + \|\mathbf{v} + \nabla q\|_{0,\Omega}^2 \gtrsim& \frac{1}{2}\|\text{div } \mathbf{v}\|_{0,\Omega}^2 + \max\left(0, 1-2C_\Omega\right)\|\nabla q\|_{0,\Omega}^2 + \min\left(1, \frac{1}{2C_\Omega}\right)\|\mathbf{v}\|_{0,\Omega}^2\\ \gtrsim& \|\text{div } \mathbf{v}\|_{0,\Omega}^2 + \|\nabla q\|_{0,\Omega}^2 + \|\mathbf{v}\|_{0,\Omega}^2\\ \gtrsim& \|\mathbf{v}\|_{\text{div},\Omega} + \|q\|_{1,\Omega}\ ,\end{aligned}$$

where in the last inequality, the Poincaré inequality (1.25) is used again. Consequently the energy balance (1.3) holds for the Poisson problem (1.33). Now it remains to construct practical subspaces of $H^1(\Omega)$ and $H(\text{div},\Omega)$. The finite element spaces that are therefore used in this work are presented in the next section.

1.5.1 The Lagrange element

The Lagrange element on an element $T \in \mathcal{T}_h$ consists on scalar-valued polynomials of degree $k \geq 1$ and the degrees of freedom are point values. As it is used to approximate $H^1(\Omega)$, then the points have to be chosen in order to ensure interelement continuity. Therefore, $k+1$ points have to be placed on each edge of T. In order to use the general interpolation theory, the standard choice is the usage of equidistant points, i.e.

$$\mathcal{N}_{E,T} = \{\mathbf{t}_j + \frac{\lambda}{k}(\mathbf{t}_i - \mathbf{t}_j) : 1 \leq i,j \leq 3,\ j \neq i, \lambda \in \mathbb{N}_0, \lambda \leq k\}\ , \tag{1.42}$$

where $\mathbf{t}_i,\ i = 1,2,3$ denote the vertices of T. Due to the fact that

$$\dim\ \mathcal{P}_k(T) = \frac{1}{2}(k+1)(k+2)\ , \tag{1.43}$$

it remains to place $\frac{1}{2}(k-2)(k-1)$ nodes on the interior of T for $k \geq 2$. This can be done by considering for each edge of T the $k+1$ parallel lines through the points on the other edge, see figure 1.1 and in particular for $k = 1$ and $k = 2$ figure 1.2.

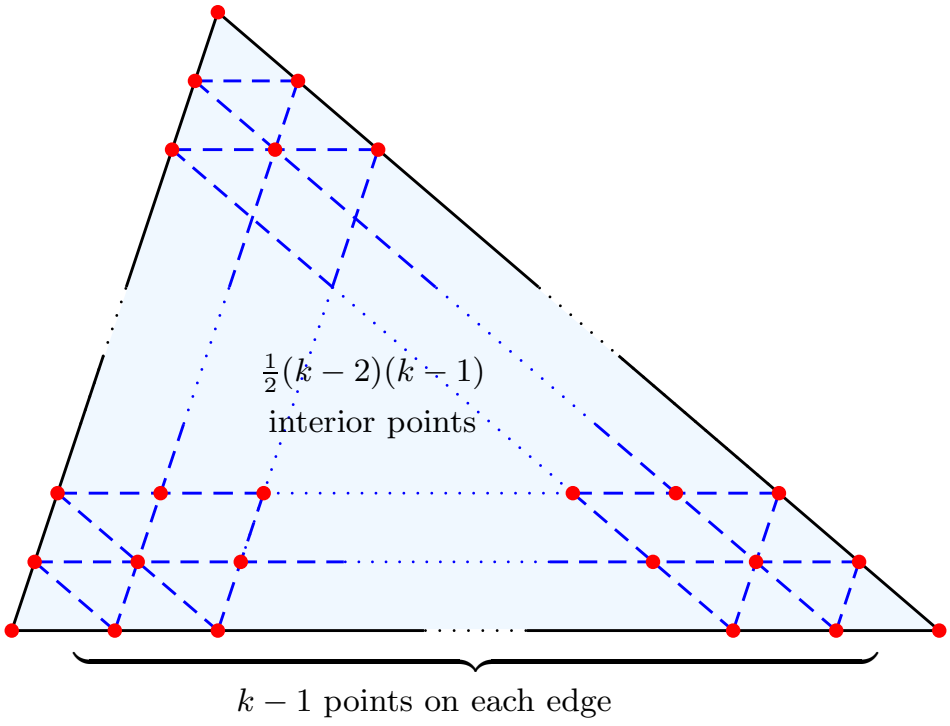


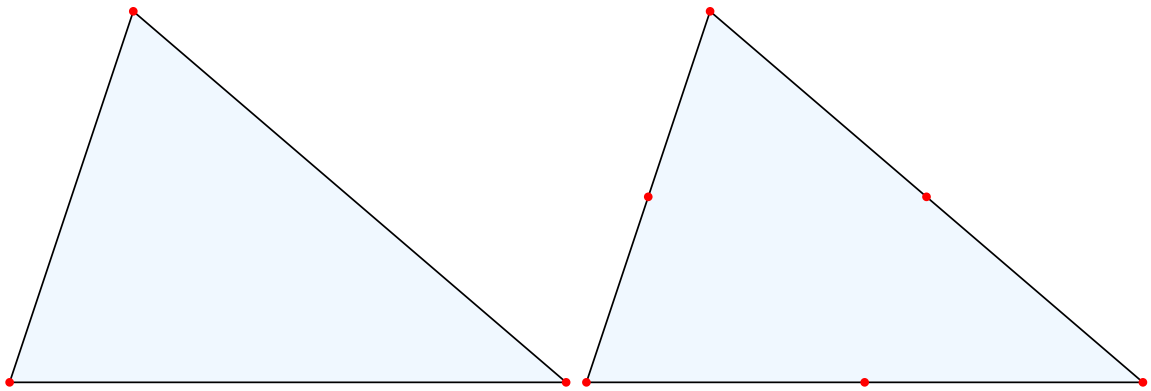
Figure 1.1: Degrees of freedom for the Lagrange element

This means for the whole set of nodal variables:

$$\mathcal{N}_T = \{\mathbf{t}_i + \sum_{j=1}^{2} \frac{\lambda_j}{k}(\mathbf{t}_j - \mathbf{t}_i) : 1 \leq i \leq 3,\ j \neq i, \lambda_j \in \mathbb{N}_0, \lambda_1 + \lambda_2 \leq k\}\ . \tag{1.44}$$

$(T, \mathcal{P}_k(T), \mathcal{N}_T)$ denotes then the Lagrange finite-element of type k. The Lagrange finite-element space $\mathcal{P}_k(\Omega, \mathcal{T}_h)$ is then (see for example [24]):

$$\mathcal{P}_k(\Omega) = \mathcal{P}_k(\Omega, \mathcal{T}_h) = \{q_h \in C^0(\overline{\Omega}) : q_{h|_T} \in \mathcal{P}_k(T)\ \forall T \in \mathcal{T}_h\} \tag{1.45}$$



(a) $k = 1$ (b) $k = 2$

Figure 1.2: Degrees of freedom for the Lagrange element for k=1,2

Note that the additional constraint $(q_h, 1)_{0,\Omega} = 0$ can be implemented in this space:

$$\dot{\mathcal{P}}_k(\Omega) = \{q_h \in C^0(\overline{\Omega}) : q_{h|_T} \in \mathcal{P}_k(T)\ \forall T \in \mathcal{T}_h \text{ and } (q_h, 1)_{0,\Omega} = 0\}\ . \tag{1.46}$$

Let $\mathcal{I}_h$ denote the corresponding interpolation operator. For each $q \in C^0(\overline{\Omega})$, it holds $\mathcal{I}_h q \in C^0(\overline{\Omega})$. Moreover, $\mathcal{I}_h q \in W^{k+1}_\infty$ and it holds for $0 \le s \le k+1$ (see theorem 4.4.4 in [14]):

$$\|q - \mathcal{I}_h q\|_{W^s_p(\Omega)} \le h^{k+1-s} |q|_{W^{k+1}_p(\Omega)}\ \forall\ q \in W^{k+1}_p(\Omega)\ . \tag{1.47}$$

In order to obtain this estimate, a key point is the scaling argument: use a reference element $(T_{\text{ref}}, \mathcal{P}_{T_{\text{ref}}}, \mathcal{N}_{T_{ref}})$ and construct the finite-elements $(T, \mathcal{P}_T, \mathcal{N}_T)$ for $T \in \mathcal{T}_h$ such that there exists an invertible affine mapping $F_{T,\text{ref}} : \mathbb{R}^2 \to \mathbb{R}^2$ with

$$T = F_{T,\text{ref}}(T_{\text{ref}}) \tag{1.48a}$$

$$\mathcal{P}_T = \{q \circ F^{-1}_{T,\text{ref}} : q \in \mathcal{P}_{T_{\text{ref}}}\} \tag{1.48b}$$

$$\mathcal{N}_T = \{l_i \circ F^{-1}_{T,\text{ref}} : l_i \in \mathcal{N}_{T_{ref}}, 1 \le i \le N\} \tag{1.48c}$$

as in figure 1.3.
Let $F_{T,\text{ref}}$ be of the form $\mathbf{x} \mapsto \alpha + B\mathbf{x}$ with a scalar α and a matrix B. Then, with the use of the Bramble-Hilbert lemma (for more details see e.g. [12]) the interpolation error can be bounded with $\|B\|$, and it holds $\|B\| \lesssim h_T$. In particular the following theorem is used with $B = J_{F_{T,\text{ref}}}$.

Theorem 1.6. *Let K and $\hat{K}$ be two bounded open subsets of $\mathbb{R}^2$ such that $K = F(\hat{K})$, where F is a one-to-one mapping that belongs to $W^1_\infty(\hat{K})$ and with inverse $F^{-1} \in W^1_\infty(K)$. Then if a function $\hat{q} : \hat{K} \to \mathbb{R}$ belongs to the space $H^1(\hat{K})$, the function $q = \hat{q} \circ F^{-1}$ belongs to the space $H^1(K)$ and it holds*

$$\nabla q = J_F^{-\top} \left((\nabla \hat{q}) \circ F^{-1} \right) \tag{1.49}$$

Moreover,

$$\|q\|^2_{0,K} \lesssim \|\det(J_F)\|_{0,\infty,\hat{K}} \|\hat{q}\|^2_{0,\hat{K}} \tag{1.50}$$

$$\|q\|^2_{1,K} \lesssim \|\det(J_F)\|_{0,\infty,\hat{K}} \|F^{-1}\|^2_{1,\infty,K} \|\hat{q}\|^2_{1,\hat{K}}\ . \tag{1.51}$$

Proof: See theorem 4.3.2. in [18].

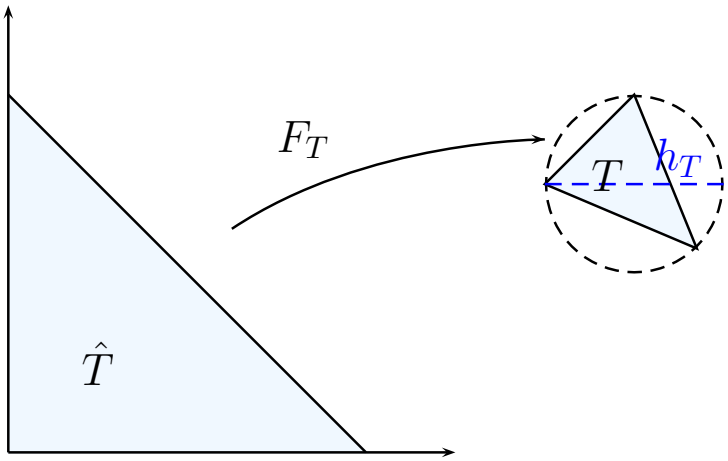


Figure 1.3: Mapping $F_{T,\text{ref}}$

1.5.2 The Raviart-Thomas Element

Contrary to the scalar-valued Lagrange elements, the Raviart-Thomas elements are vector valued and for $k \geq 0$, the space $RT_k(T)$ of the Raviart-Thomas functions on T for $T \in \mathcal{T}_h$ is a subspace of $(\mathcal{P}_{k+1}(T))^2$. The Raviart-Thomas finite element space $RT_k(\Omega) = RT_k(\Omega, \mathcal{T}_h)$ is then a subspace of the space of piecewise polynomials with respect to $\mathcal{T}_h$ in each dimension:

$$RT_k(\Omega, \mathcal{T}_h) \subset \{\mathbf{v}_h \in (L^2(\Omega))^2 : \mathbf{v}_h|_T \in RT_k(T) \forall T \in \mathcal{T}_h\} \ . \tag{1.52}$$

A sufficient condition for $H(\text{div}, \Omega)$-conformity is the normal component being continuous across each edge, see [11]. For $T \in \mathcal{T}$, let ∂T_h denote its boundary, $\mathcal{E}(T)$ the set of its edges, $\mathbf{n}$ the outward oriented normal and

$$R_k(\partial T) = \{\phi \in L^2(\partial T) : \phi|_e \in \mathcal{P}_k(e) \quad \forall \ e \in \mathcal{E}(T)\} \tag{1.53}$$

the polynomial space on the edges. The definition of the Raviart-Thomas element

$$RT_k(T) = (\mathcal{P}_k(T))^2 + \mathbf{x}\mathcal{P}_k(T) \tag{1.54}$$

ensures that some degree of freedom can be defined as the fluxes across edges of the mesh, such that $H(\text{div}, \Omega)$-conformity holds. For $\mathbf{v}_h \in RT_k(T)$, it holds

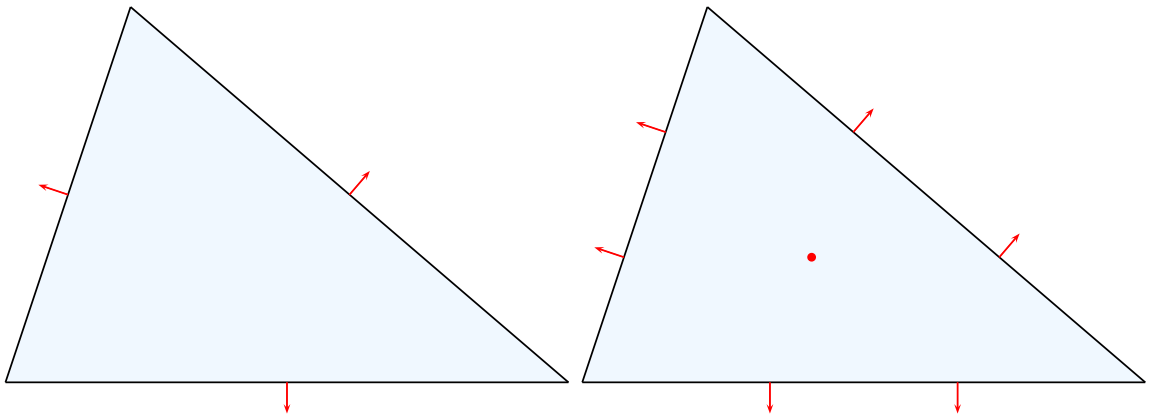
$$\begin{aligned} \text{div } \mathbf{v}_h &\in \mathcal{P}_k(T) \ , \\ \mathbf{v}_h \cdot \mathbf{n} &\in R_k(\partial T) \end{aligned} \tag{1.55}$$

and the degrees of freedom $\mathcal{N}_{RT,T}$ are given by

$$\int_{\partial T} (\mathbf{n} \cdot \mathbf{v}_h) p_k \ ds, \quad p_k \in R_k(\partial T) \ , \tag{1.56a}$$

$$\int_T \mathbf{v}_h \cdot \mathbf{p}_{k-1} \ d\mathbf{x}, \quad \mathbf{p}_{k-1} \in (P_{k-1}(T))^2 \tag{1.56b}$$

as represented in figure 1.4 for $k = 0$ and $k = 1$ (see [21] as well).



(a) $k = 0$ (b) $k = 1$

Figure 1.4: Degrees of freedom for the Raviart-Thomas element for $k = 0$ and $k = 1$

For $k \geq 1$, the points for (1.56b) can be chosen as the interior points of the Lagrange element of type $k + 2$, each point represents 2 degrees of freedom (see figure 1.5), such that

$$\dim\, RT_k(K) = (k + 3)(k + 1). \tag{1.57}$$

The proofs of the unisolvence (see [11]) allows the following definition

$$RT_k(\Omega, \mathcal{T}_h) = \{\mathbf{v}_h \in H(\mathrm{div}, \Omega) : \mathbf{v}_h|_T \in RT_k(T)\ \forall T \in \mathcal{T}_h\} \tag{1.58}$$

with the degree of freedoms described above. In [11], more details of the smoothness assumption to obtain the following interpolation theorem are given as well.

Theorem 1.7. *Let $r > 2$ be fixed, and $1 \leq m \leq k + 1$. For $\mathbf{v} \in (L^r(T))^2 \cap H^m(\Omega)$ with* div $\mathbf{v} \in L^2(\Omega)$. *Then, the degrees of freedom* (1.56a) *define an interpolation operator $\hat{\mathcal{R}}_h$, and it holds for $s \leq k + 1$:*

$$\|\mathbf{v} - \hat{\mathcal{R}}_h\mathbf{v}\|_{0,\Omega} \lesssim h^m|\mathbf{v}|_{m,\Omega} \tag{1.59a}$$

$$\|\mathrm{div}(\mathbf{v} - \hat{\mathcal{R}}_h\mathbf{v})\|_{0,\Omega} \lesssim h^s|\mathrm{div}\ \mathbf{v}|_{s,\Omega}\ . \tag{1.59b}$$

To obtain this estimate, the same idea for the scaling argument as in section 1.5.1 is used, but the affine mapping $F_{T,ref}(\mathbf{x}) = \alpha + B\mathbf{x}$ that maps $\hat{T}$ on T does not preserve the normal components, and transformed elements does not belong $H(\mathrm{div}, T)$. Instead, the Piola transformation from the following theorem has to be used (see [40]):

Theorem 1.8. *Let K and $\hat{K}$ be two bounded open subsets of $\mathbf{R}^2$ such that $K = F(\hat{K})$, where F is a one-to-one mapping that belongs to $W^1_\infty(\hat{K})$ and with inverse $F^{-1} \in W^1_\infty(K)$. Then if a function $\hat{\mathbf{v}} : \hat{K} \to \mathbf{R}^2$ belongs to the space $H(\mathrm{div}, \hat{K})$, the function*

$$\mathbf{v} = \mathcal{G}_{Piola}(\hat{\mathbf{v}}) = \left(\frac{1}{\det(J_F)} J_F\hat{\mathbf{v}}\right) \circ F^{-1} \tag{1.60}$$

belongs to the space $H(\mathrm{div}, K)$ and it holds

$$\mathrm{div}\ \mathbf{v} = \frac{\mathrm{div}\ \hat{\mathbf{v}}}{\det J_F} \circ F^{-1}\ . \tag{1.61}$$

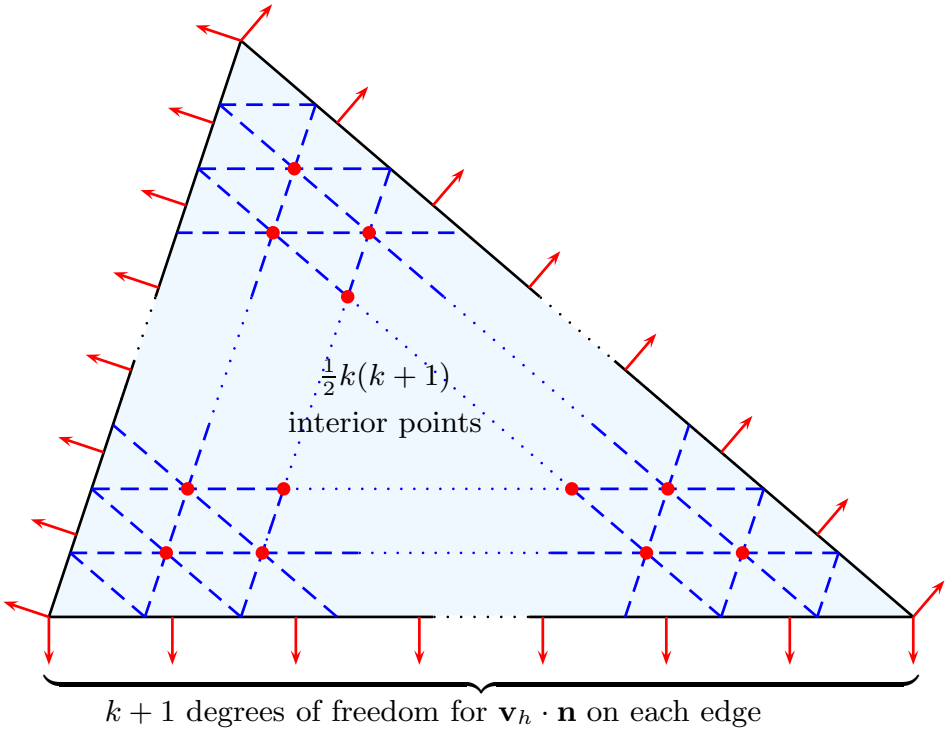


Figure 1.5: Degrees of freedom for the Raviart-Thomas element

Moreover with $q = \hat{q} \circ F^{-1}$*, it holds*

$$\int_K \mathbf{v} \cdot \nabla q \; dx = \int_{\hat{K}} \hat{\mathbf{v}} \cdot \nabla \hat{q} \; d\hat{x} \tag{1.62a}$$

$$\int_K \operatorname{div} \mathbf{v} \cdot q \; dx = \int_{\hat{K}} \operatorname{div} \hat{\mathbf{v}} \cdot \hat{q} \; d\hat{x} \tag{1.62b}$$

$$\int_{\partial K} (\mathbf{v} \cdot \mathbf{n}) q \; dx = \int_{\partial \hat{K}} (\hat{\mathbf{v}} \cdot \mathbf{n}) \hat{q} \; d\hat{x}, \tag{1.62c}$$

where $\mathbf{n}$ *denotes the normal vector of the considered element.*

Proof: see [11] lemma 2.1.6.

The remaining steps to obtain the estimates (1.59a) and (1.59b) are quite the same as in section 1.5.1, because in the affine case, the determinant has no impact on the estimates, and so the properties of B can be used again, see [35].

As the Neumann boundary conditions has to be considered, note that if $\partial_s \hat{K}$ belongs to the boundary of $\hat{K}$, the theorem 1.8 implies that $\mathcal{G}_{Piola}$ maps $H_{\partial_s \hat{K}}(\operatorname{div}, \hat{K})$ on $H_{F(\partial_s \hat{K})}(\operatorname{div}, K)$ as well. Further, due to the choice of degrees of freedom, the Neumann boundary conditions can be implemented directly in the space such that

$$RT_{k,\Gamma}(\Omega, \mathcal{T}_h) = \{\mathbf{v}_h \in RT_k(\Omega, \mathcal{T}_h) : \mathbf{v}_h|_\Gamma = 0\} \tag{1.63}$$

is a subspace of $H_\Gamma(\operatorname{div}, \Omega)$.

1.5.3 Discrete Variational Formulation

With the finite element spaces defined in the last two subsections, the discrete variational formulation of the Poisson problem is

$$\begin{aligned} &\text{Find } (\mathbf{u}_h, p_h) \in RT_{k,\Gamma}(\Omega) \times \dot{\mathcal{P}}_k(\Omega) \text{ such that} \\ &\qquad (\text{div } \mathbf{u}_h, \text{div } \mathbf{v}_h)_{0,\Omega} + (\mathbf{u}_h + \nabla p_h, \mathbf{v} + \nabla q_h)_{0,\Omega} = (f, \text{div } \mathbf{v})_{0,\Omega} \end{aligned} \tag{1.64}$$

für alle $(\mathbf{v}_h, q_h) \in RT_{k,\Gamma}(\Omega) \times \dot{\mathcal{P}}_k(\Omega)$. The solution minimizes the associated least squares functional

$$\mathcal{F}(\mathbf{u}_h, p_h) = \|\text{div } \mathbf{u}_h - f\|^2_{0,\Omega} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\Omega}, \tag{1.65}$$

under all $(\mathbf{u}_h, p_h) \in RT_{k,\Gamma}(\Omega) \times \dot{\mathcal{P}}_k(\Omega)$. Combining the theory of the section 1.2 with the properties of the interpolation operators defined in the last sections leads to the following theorem

Theorem 1.9. *Let $(\mathbf{u}, p) \in H_\Gamma(\text{div}, \Omega) \times \dot{H}^1(\Omega)$ be the exact solution of the system (1.33) and assume that it satisfies $p \in H^2(\Omega)$. Then, if $(\mathbf{u}_h, p_h) \in RT_{k,\Gamma}(\Omega) \times \dot{\mathcal{P}}_k(\Omega)$, denotes its finite element approximation,*

$$\|\mathbf{u} - \mathbf{u}_h\|_{\text{div},\Omega} + \|p - p_h\|_{1,\Omega} \lesssim h \left(\|\mathbf{u}\|_{1,\Omega} + \|f\|_{0,\Omega} + \|p\|_{2,\Omega}\right) \tag{1.66}$$

holds.

Chapter 2

Finite Elements on Domains with Curved Boundaries

From now on, $\Omega \subset \mathbb{R}^2$ is a bounded domain with a Lipschitz continuous and piecewise C^2 boundary $\Gamma = \partial\Omega$. The first section of this chapter describes the construction of a triangulation for such a curved domain. In particular, its boundary has to be interpolated, and an approximated domain $\hat{\Omega}$ with polygonal boundary $\hat{\Gamma}$ is defined.

Further, the finite-element spaces on Ω are introduced in the second section. In particular, the space $RT_{0,\hat{\Gamma}}(\Omega)$ is defined to approximate $H_\Gamma(\mathrm{div}, \Omega)$.

The third section presents the construction of a mapping from $\hat{\Omega}$ to Ω, as done in [27] and some of its properties. This allows to connect the least squares functional on $\hat{\Omega}$ to the one on Ω in Chapter 3 and is the point of departure of the further analysis of the least squares method on domains with curved boundaries.

As the normal flux of functions that belongs to $RT_{0,\hat{\Gamma}}(\Omega)$ vanishes only on $\hat{\Gamma}$ and does not vanish on Γ, an estimate for its value on Γ is derived in the fourth section. These results will be used in Chapter 3 as well in order to bound the least squares functional from below.

2.1 Construction of the Curved Triangulation

In this section, the concept of triangulation is extended to a domain with curved boundaries. Therefore, let $[t_0, t_N] \subset \mathbb{R}$ with $t_0 < t_N$ denote the domain of Γ, i.e.

$$\Gamma = \{\boldsymbol{\gamma}(t) \ : \ t_0 \leq t \leq t_N\} \ . \tag{2.1}$$

and assume that $\boldsymbol{\gamma} : [t_0, t_N) \to \mathbb{R}^2$ is injective and that $\boldsymbol{\gamma}(t_0) = \boldsymbol{\gamma}(t_N)$. The first step to construct a triangulation for Ω is to obtain an approximated boundary by piecewise linear interpolation, in particular the points $0 < t_i < t_N, \ i = 1, ..., N-1$ such that

(i) $i < j$ implies $t_i < t_j$ for $i, j = 0, ..., N$.

(ii) all points where Γ is not C^2 are in the set of interpolation points $\{\boldsymbol{\gamma}(t_i)\}_{i=0}^N$.

(iii) the polygon $\hat{\Omega}$ that is formed with the vertex $\{\boldsymbol{\gamma}(t_i)\}_{i=0}^N$ is simple.

(iv) for each curved boundary segement $\Gamma_i = \{\boldsymbol{\gamma}(t) \ : \ t_{i-1} \leq t \leq t_i\}$, the diameter of the smallest circle that completely contains it is smaller than h.

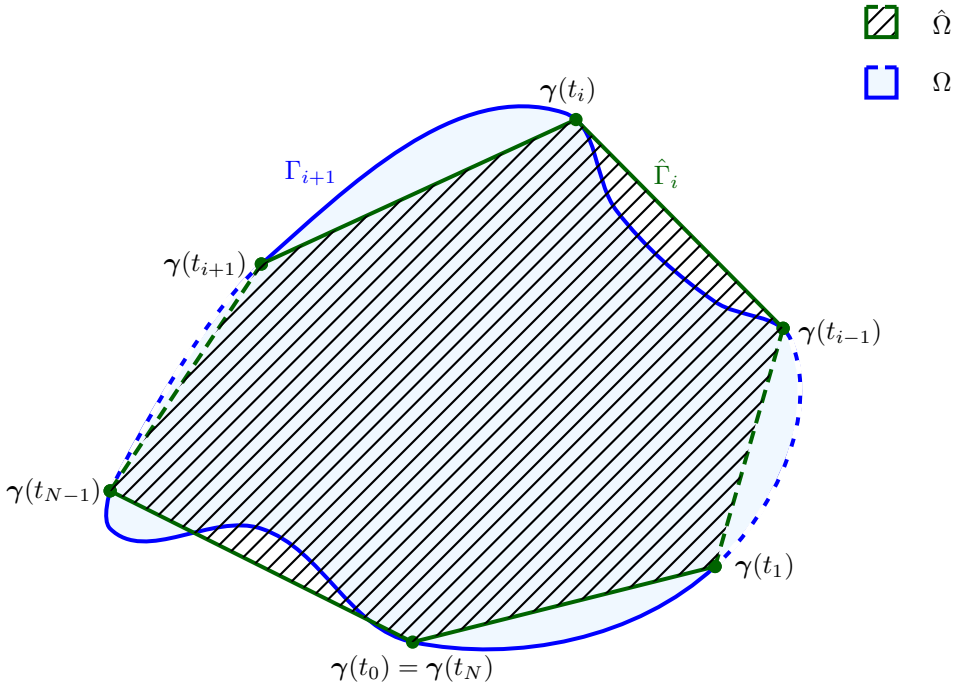


Figure 2.1: Interpolation of the boundary

Let $\hat{\Gamma}$ be the boundary of $\hat{\Omega}$ (see Figure 2.1). It can be parametrized as follows

$$\hat{\Gamma} = \bigcup_{i=1}^{N} \left\{ \hat{\boldsymbol{\gamma}}_i(t) = \frac{t_i - t}{t_i - t_{i-1}} \boldsymbol{\gamma}(t_{i-1}) + \frac{t_{i-1} - t}{t_{i-1} - t_i} \boldsymbol{\gamma}(t_i) \ : \ 1 \leq i \leq N \text{ and } t_{i-1} \leq t \leq t_i \right\} \tag{2.2}$$

and the length of each boundary segement $\hat{\Gamma}_i = \{\hat{\boldsymbol{\gamma}}(t) \ : \ t_{i-1} \leq t \leq t_i\}$ is smaller as h due to the fourth property above. Hence, note that from the classical interpolation theory it holds

$$\|\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}}\|_{l,\infty,[t_0,t_N]} \lesssim h^2 |D^2 \boldsymbol{\gamma}|_{\infty,[t_0,t_N]} \qquad \forall l \leq 2 \tag{2.3}$$

and therefore

$$|\Omega \backslash \Omega_h| \eqsim |\Omega_h \backslash \Omega| \eqsim h^2. \tag{2.4}$$

Now, let $\hat{\mathcal{T}}_h$ be a quasi-uniform triangulation of $\hat{\Omega}$ as described in Section 1.2, which consists completely of straight triangles. In order to restrict the analysis on a single element before summing over all elements, the following notation is introduced:

(i) Assume that the triangles $\hat{T}$ in the triangulation $\hat{\mathcal{T}}$ are numbered such that for $i \leq N$, $\hat{T}_i$ is the triangle for which $\hat{T}_i \cap \hat{\Gamma}_i \neq 0$. For $N < i \leq \bar{N}$, $\hat{T}_i$ is located in the interior of the domain.

(ii) For $i \leq N$, $\mathscr{V}_3(\hat{T}_i)$ denotes the vertex of $\hat{T}_i$ that does not belong to $\hat{\Gamma}_i$.

(iii) $\widetilde{T}_i$ is the curved triangle corresponding to $\hat{T}_i$. This is a shape with the same vertices as $\hat{T}_i$ and whose boundary consists of the segments $\partial \hat{T}_i \backslash \hat{\Gamma}_i$ and the arc Γ_i.

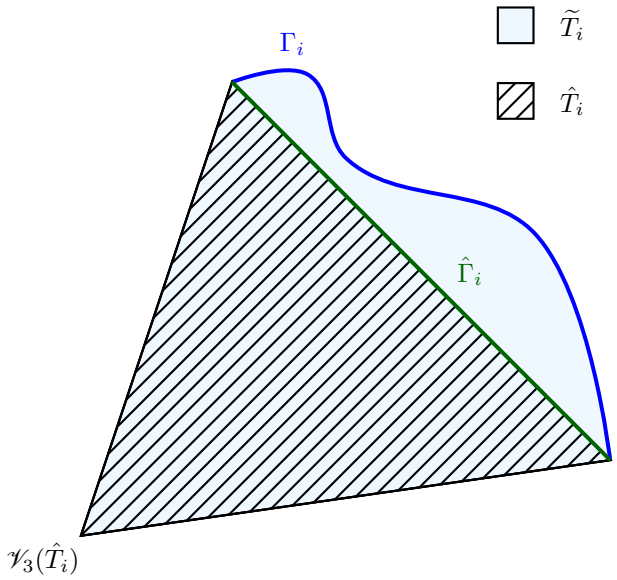


Figure 2.2: Notations for a curved triangle

(iv) h_i denotes the length of $\hat{\Gamma}_i$.

These notations are summarized in Figure 2.2. Now, it is possible to define the curved triangulation

$$\widetilde{\mathcal{T}}_h = \{\hat{T}_i \ : \ i > N\} \cup \{\widetilde{T}_i \ : \ 1 \leq i \leq N\} \tag{2.5}$$

such that Ω is completely covered, as illustrated in Figure 2.3

2.2 Finite Elements on the Curved Triangulation

Now, for an arbitrary finite-element space $\mathcal{V}_h$ defined with respect to $\hat{\mathcal{T}}_h$, the piecewise polynomial functions $v_h \in \mathcal{V}_h$ may be extended to $\widetilde{T} \in \widetilde{\mathcal{T}}_h$ by using its polynomial representation on the corresponding element $\hat{T} \in \hat{\mathcal{T}}_h$. Note that as for $i \neq j$ the symmetric difference $\widetilde{\Delta}_i = \widetilde{T}_i \backslash \hat{T}_i \cup \hat{T}_i \backslash \widetilde{T}_i$ has no common point with $\widetilde{\Delta}_j$, this has no impact on the smoothness of the finite-element functions. For the finite-element spaces used to solve the Poisson problem, this means

$$\mathcal{P}_k(\Omega) = \mathcal{P}_k(\Omega, \widetilde{\mathcal{T}}_h) = \{q_h \in C^0(\overline{\Omega}) : q_{h|\widetilde{T}} \in \mathcal{P}_k(\widetilde{T}) \ \forall \widetilde{T} \in \widetilde{\mathcal{T}}_h\} \tag{2.6}$$

and

$$RT_k(\Omega, \widetilde{\mathcal{T}}_h) = RT_k(\Omega, \widetilde{\mathcal{T}}_h) = \{\mathbf{v}_h \in H(\mathrm{div}, \Omega) : \mathbf{v}_h|_{\widetilde{T}} \in RT_k(\widetilde{T}) \ \forall \widetilde{T} \in \widetilde{\mathcal{T}}_h\}. \tag{2.7}$$

For these Raviart-Thomas elements, the Neumann boundary conditions have to be set on $\hat{\Gamma}$, with the consequence that the resulting space is not a subspace of $H_\Gamma(\mathrm{div}, \Omega)$. In Section 2.4, an estimate for $\mathbf{v}_h \cdot \mathbf{n}$ on Γ will be given for Raviart-Thomas elements of lowest order $\mathbf{v}_h$ with boundary conditions $\mathbf{n} \cdot \mathbf{v}_h = 0$ on $\hat{\Gamma}$, i.e. for

$$\mathbf{v}_h \in RT_{0,\hat{\Gamma}}(\Omega) = \{\mathbf{v}_h \in H_{\hat{\Gamma}}(\mathrm{div}, \Omega) : \mathbf{v}_h|_{\widetilde{T}}(\mathbf{x}) = \boldsymbol{\alpha} + \beta \mathbf{x}, \boldsymbol{\alpha} \in \mathbb{R}^2, \beta \in \mathbb{R} \quad \forall \widetilde{T} \in \widetilde{\mathcal{T}}_h\}. \tag{2.8}$$

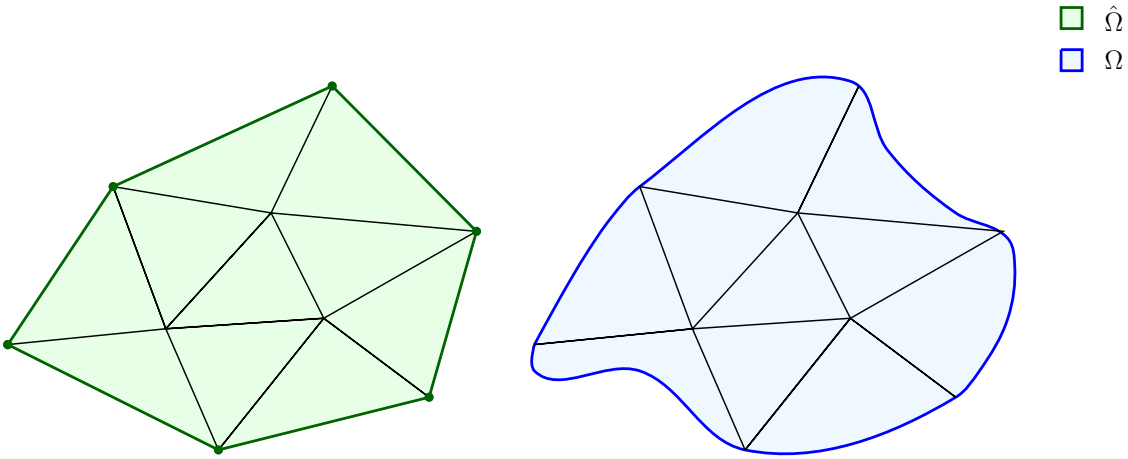


Figure 2.3: Triangulation of $\hat{\Omega}$ and Ω

For the Lagrange elements with the additional condition $(q_h, 1)_{0,\Omega} = 0$, the approximation space corresponding to the Raviart-Thomas elements lowest order case is

$$\acute{\mathcal{P}}_1(\Omega) = \{q_h \in C^0(\overline{\Omega}) : q_{h|\widetilde{T}} \in \mathcal{P}_1(\widetilde{T}) \ \forall \widetilde{T} \in \widetilde{\mathcal{T}}_h, \text{ and } (q_h, 1)_{0,\hat{\Omega}} = 0\} \tag{2.9}$$

where the additional condition is replaced with an implementable approximated one. However, in order to simplify the analysis in the next chapter, the following space is used for the analysis

$$\dot{\mathcal{P}}_1(\Omega) = \{q_h \in C^0(\overline{\Omega}) : q_{h|\widetilde{T}} \in \mathcal{P}_1(\widetilde{T}) \ \forall \widetilde{T} \in \widetilde{\mathcal{T}}_h, \text{ and } (q_h, 1)_{0,\Omega} = 0\} \tag{2.10}$$

since the two solutions $\acute{p}_h \in \acute{\mathcal{P}}_1(\Omega)$ and $p_h \in \dot{\mathcal{P}}_1(\Omega)$ only differ by a constant on the order of h. This will be proven in Section 2.5 (see Lemma 2.4), as the estimate about the normal flux on interpolated boundaries from the Section 2.4 is first needed.

2.3 Mapping from $\hat{\Omega}$ to Ω

In order to simplify the notations for the construction of the mapping $\hat{\mathbf{\Phi}}_h$ from $\hat{\Omega}$ to Ω, a reference element is needed. In this work, the reference triangle is chosen as the right triangle with the vertices $(0,0)$, $(1,0)$, $(0,1)$, and the transformation mapping it to the triangle $\hat{T}_i \in \hat{\mathcal{T}}_h$ and $E_{\text{ref}} = [0,1]$ to $\hat{\Gamma}_i$ is denoted with $\mathbf{F}_{\text{ref},\hat{T}_i}$.

On this reference triangle, a mapping $\mathbf{Z}_{\text{ref}}$ to connect an interior point with a corresponding point on the edge E_{ref} is needed. Therefore, for a point $\mathbf{x}_{\text{ref}} = (x_{\text{ref}}, y_{\text{ref}}) \in \hat{T}_{\text{ref}}$ consider the line outgoing from $(0,1)$ throught the point $\mathbf{x}_{\text{ref}}$. As it crosses the edge $[0,1]$ at the point $(\frac{x_{\text{ref}}}{1-y_{\text{ref}}}, 0)$ the mapping $\mathbf{Z}_{\text{ref}}$ is defined as follows.

$$\mathbf{Z}_{\text{ref}} : \hat{T}_{\text{ref}} \to \quad [0,1] \tag{2.11}$$

$$(x_{\text{ref}}, y_{\text{ref}}) \mapsto \quad \frac{x_{\text{ref}}}{1 - y_{\text{ref}}}. \tag{2.12}$$

With this mapping $\mathbf{Z}_{\text{ref}}$, a mapping $\mathbf{Z}_i = \mathbf{F}_{\text{ref},\hat{T}_i} \circ \mathbf{Z}_{\text{ref}} \circ \mathbf{F}^{-1}_{\text{ref},\hat{T}_i}$ can be defined to connect an interior point of any triangle $\hat{T}_i$ (whose intersection with $\hat{\Gamma}$ is not empty) with a corresponding

point on $\hat{\Gamma}_i$, as illustrated in Figure 2.4. Due to the construction of this mapping, any point $\mathbf{x} \in \hat{T}_i$ is located on the line segment $[\mathscr{V}_3(\hat{T}_i), \mathbf{Z}_i(\mathbf{x})]$. The position of $\mathbf{x}$ on this line can then be given with the ratio $\delta(\mathbf{x})$ of the distance between $\mathscr{V}_3(\hat{T}_i)$ and $\mathbf{x}$ to the distance between $\mathscr{V}_3(\hat{T}_i)$ and $\mathbf{Z}_i(\mathbf{x})$, i.e :

$$\delta(\mathbf{x}) = \frac{\text{dist}(\mathscr{V}_3(\hat{T}_i), \mathbf{x})}{\text{dist}(\mathscr{V}_3(\hat{T}_i), \mathbf{Z}_i(\mathbf{x}))} \tag{2.13}$$

Note that due to the fact that the affine mapping conserves the ratio of the distances, it holds

$$\delta(\mathbf{x}) = \frac{\text{dist}\left((0,1), \mathbf{F}^{-1}_{\text{ref},\hat{T}_i}(\mathbf{x})\right)}{\text{dist}\left((0,1), \mathbf{Z}_{\text{ref}}(\mathbf{F}^{-1}_{\text{ref},\hat{T}_i}(\mathbf{x}))\right)}. \tag{2.14}$$

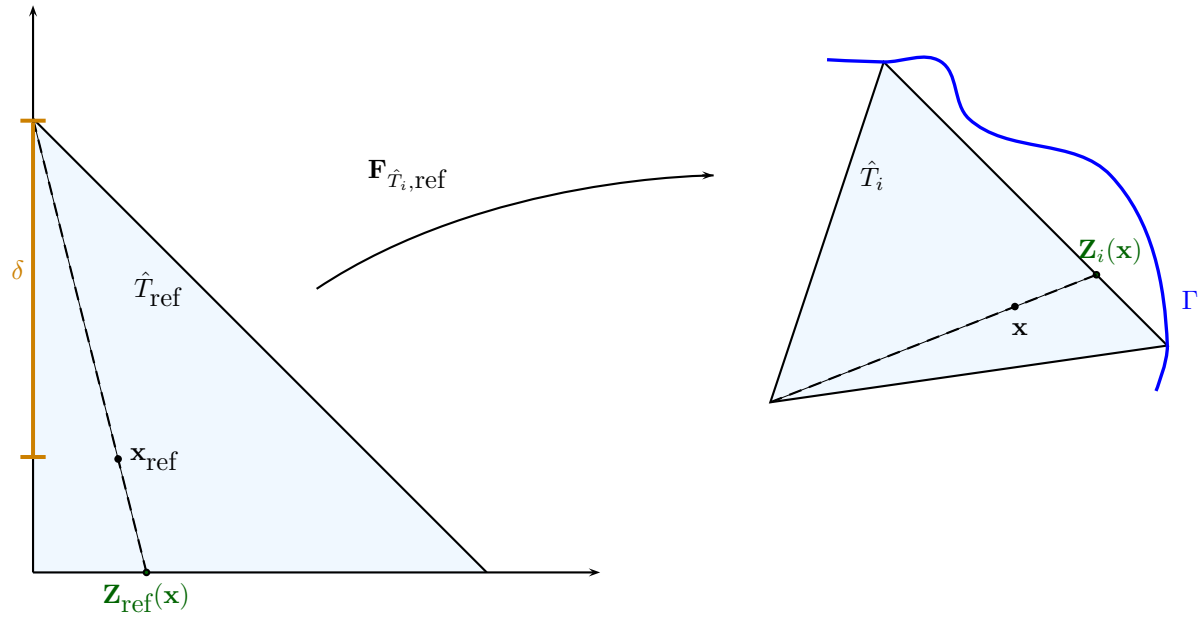


Figure 2.4: Construction of the mapping $\mathbf{Z}_i$

An example of the streching that results from $\mathbf{Z}_i$ is represented in Figure 2.5. The next step in the construction of $\hat{\mathbf{\Phi}}_h$ is to connect $\mathbf{Z}_i(\mathbf{x})$ into a point on Γ_i. Of course, $\boldsymbol{\gamma}$ has to be used, but therefore, the point $\mathbf{Z}_i(\mathbf{x})$ has first to be mapped back in the domain of the chart. This can be done with $\hat{\boldsymbol{\gamma}}$ as it is invertible due to its construction. Therefore, define $\hat{\boldsymbol{\zeta}} = \boldsymbol{\gamma} \circ \hat{\boldsymbol{\gamma}}^{-1}$, such that $\hat{\boldsymbol{\zeta}}$ maps $\hat{\Gamma}$ on Γ. Then, the point $\hat{\boldsymbol{\zeta}}(\mathbf{Z}(\mathbf{x}))$ has to be mapped on the interior on the triangle. As the mapping $\hat{\mathbf{\Phi}}_{h,i} = \hat{\mathbf{\Phi}}_{h|_{\hat{T}_i}}$ has to be the identity map on the edges of $\hat{T}_i$ that are not mapped on the curved edge, the point $\hat{\boldsymbol{\zeta}}(\mathbf{z}(\mathbf{x}))$ has to be mapped back on the line through the third point of the triangle, such that the ratio of the distance between $\mathscr{V}_3(\hat{T}_i)$ and $\hat{\boldsymbol{\zeta}}(\mathbf{Z}_i(\mathbf{x}))$ to the distance between and $\mathscr{V}_3(\hat{T}_i)$ and the mapped back point is equal to $\delta(\mathbf{x})$. This is illustrated in Figure 2.6.

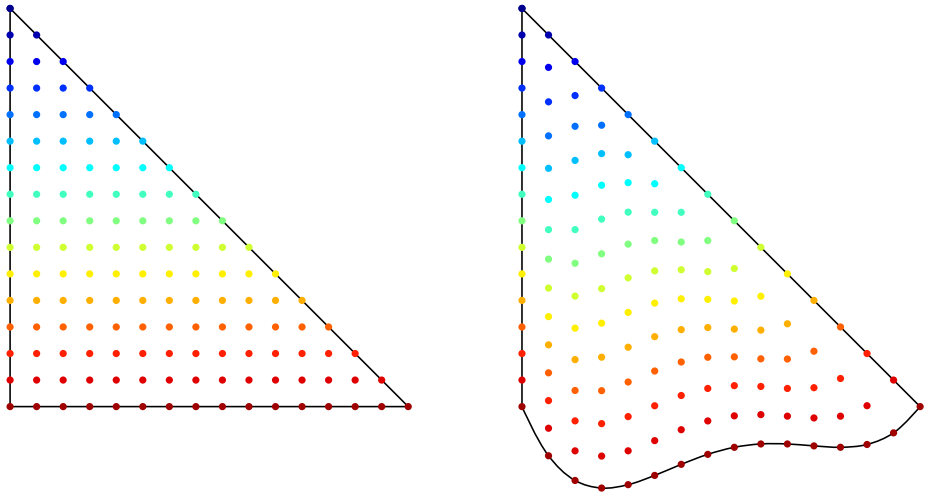


Figure 2.5: Example of the streching resulting from $\mathbf{Z}_i$

Let $\mathbf{Y}_{i,\delta} : \Gamma_i \to \widetilde{T}_i$ denote this mapping:

$$\mathbf{Y}_{i,\delta} = \mathbf{F}^{-1}_{\text{ref},\hat{T}_i} \circ \mathbf{Y}_{\text{ref},\delta} \circ \mathbf{F}_{\text{ref},\hat{T}_i} \tag{2.15}$$

where

$$\begin{aligned} \mathbf{Y}_{\text{ref},\delta} : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto (\delta x, \delta y - \delta + 1) \end{aligned} \tag{2.16}$$

Now, the mapping $\hat{\mathbf{\Phi}}_{h,i}$ can be defined as

$$\hat{\mathbf{\Phi}}_{h,i}(\mathbf{x}) = \mathbf{Y}_{i,\delta(\mathbf{x})}(\hat{\boldsymbol{\zeta}}(\mathbf{Z}_i(\mathbf{x}))). \tag{2.17}$$

Setting $\hat{\mathbf{\Phi}}_{h,i} = \text{id}$ for the interior elements leads to the definition of $\hat{\mathbf{\Phi}}_h$.

Now, some properties of $\hat{\mathbf{\Phi}}_h$ have to be derived. Therefore, first note that due to its construction, $\hat{\mathbf{\Phi}}_{h,i}$ is the identity mapping on the two edges of $\hat{T}_i$ each of which has one single common point with $\hat{\Gamma}$ and it holds

$$\hat{\mathbf{\Phi}}_h \circ \hat{\boldsymbol{\gamma}} = \boldsymbol{\gamma}. \tag{2.18}$$

and

$$\begin{aligned} \hat{\mathbf{\Phi}}_{h,i}(\mathbf{x}) - \mathbf{x} &= \mathbf{Y}_{\delta(\mathbf{x})}((\hat{\boldsymbol{\zeta}}(\mathbf{Z}(\mathbf{x}))) - \mathbf{Y}_{\delta(\mathbf{x})}(\mathbf{Z}(\mathbf{x}))) \\ &= \mathbf{Y}_{\delta(\mathbf{x})}((\hat{\boldsymbol{\zeta}} - \text{id})(\mathbf{Z}(\mathbf{x})))) \\ &= \mathbf{Y}_{\delta(\mathbf{x})}((\boldsymbol{\gamma} \circ \hat{\boldsymbol{\gamma}}^{-1} - \hat{\boldsymbol{\gamma}} \circ \hat{\boldsymbol{\gamma}}^{-1})(\mathbf{Z}(\mathbf{x})))) \\ &= \mathbf{Y}_{\delta(\mathbf{x})}(((\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}}) \circ \hat{\boldsymbol{\gamma}}^{-1})(\mathbf{Z}(\mathbf{x})))) \end{aligned} \tag{2.19}$$

This leads to the following Lemma (see [27] Paragraph 5)

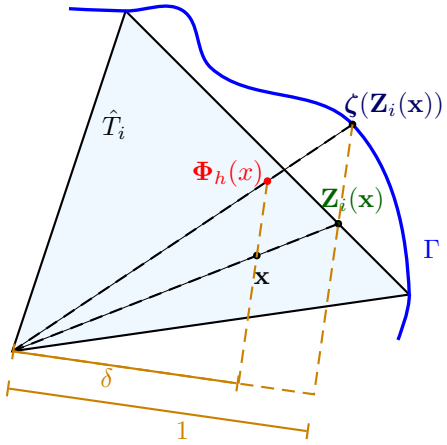


Figure 2.6: Construction of the mapping $\hat{\mathbf{\Phi}}_h$

Lemma 2.1. *With the previous assumptions and $\hat{T}_i \in \hat{\mathcal{T}}_h$, $\hat{\mathbf{\Phi}}_{h,i}$ is a $\mathcal{C}^2$ diffeomorphism and it holds*

$$\|\hat{\mathbf{\Phi}}_{h,i} - id\|_{W^j_\infty(\hat{T}_i)} \lesssim h^{2-j} \quad \forall j \leq 2 \tag{2.20}$$

$$\|\det(J_{\hat{\mathbf{\Phi}}_{h,i}}) - 1\|_{\infty,\hat{T}_i} \lesssim h \tag{2.21}$$

Moreover, $J_{\hat{\mathbf{\Phi}}_{h,i}}$ is invertible and $\hat{\mathbf{\Phi}}_{h,i} : \hat{T}_i \to \widetilde{T}_i$ is injective. The mapping $\hat{\mathbf{\Psi}}_{h,i} = \hat{\mathbf{\Phi}}^{-1}_{h,i}$ satisfies

$$\begin{aligned} \|\hat{\mathbf{\Psi}}_{h,i} - id\|_{W^j_\infty(\widetilde{T}_i)} &\lesssim h^{2-j} \quad \forall j \leq 2 \\ \|\det(J_{\hat{\mathbf{\Psi}}_{h,i}}) - 1\|_{\infty,\widetilde{T}_i} &\lesssim h \end{aligned} \tag{2.22}$$

With these properties, the following lemma needed in the construction of an approximated functional in chapter 3 can be proven.

Lemma 2.2. *Let f belong to $W^1_\infty(\hat{\Omega} \cup \widetilde{\Omega})$. Then, it holds*

$$\left|(f,1)_{0,\hat{T}_i} - (f,1)_{0,\widetilde{T}_i}\right| \lesssim h_i^3 \|f\|_{W^1_\infty(\hat{T}_i \cup \widetilde{T}_i)} + h_i^2 \|f\|_{0,\widetilde{T}_i} . \tag{2.23}$$

for $1 \leq i \leq N$.

Proof: First, note that due to the above definition of $\hat{\mathbf{\Psi}}_h$, it holds

$$(f,1)_{0,\hat{T}_i} - (f,1)_{0,\widetilde{T}_i} = \int_{\widetilde{T}_i} \left(f(\hat{\mathbf{\Psi}}_h(\mathbf{x})) \det J_{\hat{\mathbf{\Psi}}_h}(\mathbf{x}) - f(\mathbf{x}) \right) \mathbf{dx} \tag{2.24}$$

which leads to

$$\left|(f,1)_{0,\hat{T}_i} - (f,1)_{0,\widetilde{T}_i}\right| \lesssim \int_{\widetilde{T}_i} \left| f(\hat{\mathbf{\Psi}}_h(\mathbf{x})) - f(\mathbf{x}) \right| \mathbf{dx} + \|f\|_{0,\widetilde{T}_i} \|\det J_{\hat{\mathbf{\Psi}}_h} - 1\|_{0,\widetilde{T}_i} . \tag{2.25}$$

Further, for all $\mathbf{x} \in \widetilde{T}_i$, it holds

$$\begin{aligned} \left| f(\hat{\mathbf{\Psi}}_h(\mathbf{x})) - f(\mathbf{x}) \right| &= \left| \int_0^1 \nabla f(\mathbf{x} + s(\hat{\mathbf{\Psi}}_h(\mathbf{x}) - \mathbf{x})) \cdot (\hat{\mathbf{\Psi}}_h(\mathbf{x}) - \mathbf{x})\, ds \right| \\ &\leq \int_0^1 \left| \nabla f(\mathbf{x} + s(\hat{\mathbf{\Psi}}_h(\mathbf{x}) - \mathbf{x})) \right| \left| \hat{\mathbf{\Psi}}_h(\mathbf{x}) - \mathbf{x} \right| ds \end{aligned}$$

such that using (2.22), (2.25) turns into

$$\left|(f,1)_{0,\hat{T}_i} - (f,1)_{0,\widetilde{T}_i}\right| \lesssim h_i^3 \|f\|_{W^1_\infty(\hat{T}_i \cup \widetilde{T}_i)} + h_i^2 \|f\|_{0,\widetilde{T}_i} . \tag{2.26}$$

□

2.4 An Estimate for the Normal Flux on Interpolated Boundaries

As $RT_{0,\hat{\Gamma}}(\Omega)$ is not a subspace of $H_{\hat{\Gamma}}(\mathrm{div},\Omega)$, in general, for a curved boundary $\Gamma = \partial\Omega$, the normal flux $\mathbf{v}_h \cdot \mathbf{n}$ on Γ has to be estimated for $\mathbf{v}_h \in 0,\hat{\Gamma}$. Such an estimate is stated by the following theorem, which is published as Theorem 2.1 in [7].

Theorem 2.1. *Let Ω have the properties stated in Section 2.1. Then,*

$$|\langle \mathbf{n}\cdot\mathbf{v}_h, q\rangle_{0,\Gamma}| \lesssim h \left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\mathrm{div}\ \mathbf{v}_h\|^2_{0,\Omega}\right)^{\frac{1}{2}} \|q\|_{\frac{1}{2},\Gamma} \tag{2.27}$$

holds for all $\mathbf{v}_h \in RT_{0,\hat{\Gamma}}(\Omega)$ and $q \in H^{\frac{1}{2}}(\Gamma)$.

Proof: Consider the curved boundary edge Γ_i. To simplify the notation, first note that the smoothness assumption on Γ implies that the coordinate system can be shifted and rotated with a direct isometry $\mathbf{D}_i$ that mapped $\hat{\Gamma}_i$ on $[0,h_i]$, as represented in Figure 2.7. The image of the curved edge Γ_i under $\mathbf{D}_i$ can then be given as the graph of a function η:

$$\mathbf{D}_i(\Gamma_i) = \Gamma_{\mathbf{D},i} = \left\{ \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix} : 0 \le \xi \le h_i \right\} . \tag{2.28}$$

where

$$\begin{aligned} \eta : [0,h_i] &\to \mathbb{R}^2 \\ \xi &\mapsto [\mathbf{D}_i\left(\hat{\boldsymbol{\zeta}}\left(\mathbf{D}_i^{-1}\left((0,\xi)\right)\right)\right)]_2. \end{aligned} \tag{2.29}$$

As $\hat{\boldsymbol{\zeta}}$ is a C^2-diffeomorphism, η is a C^2-function. Thus, due to $\eta(0) = \eta(h_i) = 0$, there is a $\bar{\xi} \in [0,h_i]$ such that $\eta'(\bar{\xi}) = 0$. This implies

$$|\eta'(\xi)| \lesssim h_i \tag{2.30}$$

$$\text{and } |\eta(\xi)| \lesssim h_i^2 \tag{2.31}$$

for all $\xi \in [0,h_i]$.

Further, define $\mathbf{v}_{h,i} = \mathbf{v}_h|_{\widetilde{T}_i}$ and $\mathbf{w}_{h,i} = \mathbf{v}_{h,i} \circ \mathbf{D}_i^{-1}$and note that due to $\det(J_{\mathbf{D}_i}) = 1$ it holds

$$\mathrm{div}\ \mathbf{w}_{h,i} = \mathrm{div}\ \mathbf{v}_{h,i} \circ \mathbf{D}_i^{-1} \tag{2.32}$$

and

$$\int_{\Gamma_i} (\mathbf{v}_h \cdot \mathbf{n}) q \ \mathrm{d}s = \int_0^{h_i} (\mathbf{w}_{h,i} \cdot \mathbf{n})(q \circ \mathbf{D}_i^{-1}) \ \mathrm{d}\xi \tag{2.33}$$

for $q \in L^2(\hat{\Gamma}_i)$

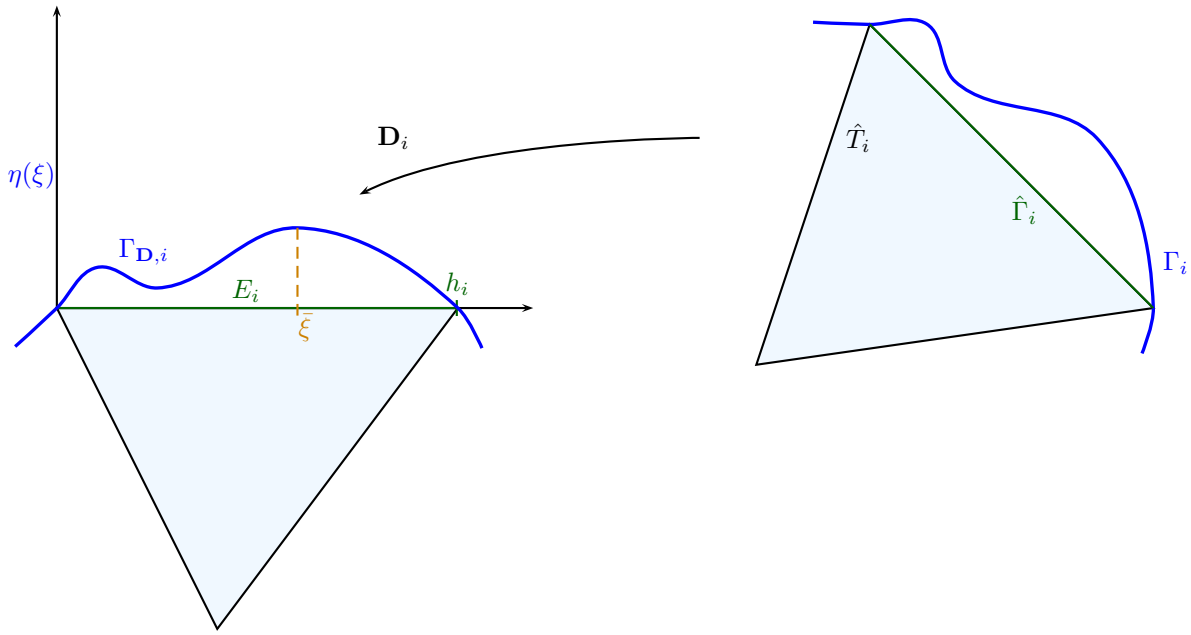


Figure 2.7: Polygonal approximation of a curved triangle

The advantage of this transformation is that the special form of $\mathbf{w}_{h,i} \in RT_{0,E_i}(\mathbf{D}_i(\widetilde{T}_i))$ can be used. The fact that $\mathbf{w}_{h,i}(\mathbf{x})\cdot\mathbf{n} = 0$ has to hold for all $\mathbf{x}$ on the edge $E_i = \{(\xi, 0) \ : \ 0 \leq \xi \leq h_i\}$ leads to

$$\mathbf{w}_{h,i}(\xi, 0) = \left(\boldsymbol{\alpha} + \beta \begin{pmatrix} \xi \\ 0 \end{pmatrix}\right) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \quad \forall \xi \in [0, h_i] \tag{2.34}$$

and thus to

$$\mathbf{w}_{h,i}(\mathbf{x}) = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \beta \mathbf{x} \quad \alpha, \beta \in \mathbb{R}. \tag{2.35}$$

This implies

$$\operatorname{div} \mathbf{w}_{h,i} \equiv 2\beta \tag{2.36}$$

$$\mathbf{w}_{h,i}(\xi, 0) = \begin{pmatrix} \alpha + \xi\, \beta \\ 0 \end{pmatrix} \text{ on } E_i \tag{2.37}$$

$$\text{and } \mathbf{n} \cdot \mathbf{w}_{h,i} = \frac{1}{(1 + \eta'(\xi)^2)^{\frac{1}{2}}} \begin{pmatrix} -\eta'(\xi) \\ 1 \end{pmatrix} \cdot \begin{pmatrix} \alpha + \xi\, \beta \\ \eta(\xi)\, \beta \end{pmatrix} \text{ on } \Gamma_{\mathbf{D},i}. \tag{2.38}$$

For $q \in L^2(\Gamma_i)$, this leads to

$$\begin{aligned}\langle \mathbf{n}\cdot\mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} &= \langle \mathbf{n}\cdot\mathbf{w}_{h,i}, q_{\mathbf{D}}\rangle_{0,\Gamma_{\mathbf{D}_i}} \\ &= \int_0^{h_i} \left(-\eta'(\xi)(\alpha+\xi\,\beta)+\eta(\xi)\,\beta\right)\, q_{\mathbf{D}}(\xi,\eta(\xi))\,d\xi \\ &\lesssim \int_0^{h_i} \left(h_i\,|\alpha+\xi\,\beta| + h_i^2|\beta|\right)|q_{\mathbf{D}}(\xi,\eta(\xi))|\,d\xi\,.\end{aligned} \tag{2.39}$$

where $q_{\mathbf{D}} = q\circ\mathbf{D}_i^{-1}$. Hence, due to (2.37) and (2.36) it holds

$$\begin{aligned}\langle \mathbf{n}\cdot\mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} &\lesssim \left(h_i\|\mathbf{w}_{h,i}\|_{0,E_i} + h_i^2\|\operatorname{div}\mathbf{w}_{h,i}\|_{0,E_i}\right)\|q_{\mathbf{D}}\|_{0,\Gamma_{\mathbf{D}_i}} \\ &\lesssim h_i\left(\|\mathbf{w}_{h,i}\|_{0,E_i}^2 + \|\operatorname{div}\mathbf{w}_{h,i}\|_{0,E_i}^2\right)^{\frac{1}{2}}\|q\|_{0,\Gamma_i} \\ &\leq h_i\left(\|\mathbf{v}_{h,i}\|_{0,\partial\hat{T}_i}^2 + \|\operatorname{div}\mathbf{v}_{h,i}\|_{0,\partial\hat{T}_i}^2\right)^{\frac{1}{2}}\|q\|_{0,\Gamma_i}\end{aligned} \tag{2.40}$$

At this point, a scaling argument has to be used to bound $\|\mathbf{v}_{h,i}\|_{0,\partial\hat{T}_i}$ by $h_i^{-\frac{1}{2}}\|\mathbf{v}_{h,i}\|_{0,\hat{T}_i}$, see for example [11]:

$$\begin{aligned}\|\mathbf{v}_{h,i}\|_{0,\hat{T}_i} &\eqsim h_i\|\mathbf{v}_{h,i}\circ\mathbf{F}_{\hat{T}_i,\mathrm{ref}}\|_{0,\hat{T}_{\mathrm{ref}}} \eqsim h_i\|\mathbf{v}_{h,i}\circ\mathbf{F}_{\hat{T}_i,\mathrm{ref}}\|_{\infty,\hat{T}_{\mathrm{ref}}} \\ &\eqsim h_i\|\mathbf{v}_{h,i}\circ\mathbf{F}_{\hat{T}_i,\mathrm{ref}}\|_{\infty,\partial\hat{T}_{\mathrm{ref}}} \eqsim h_i^{\frac{1}{2}}\|\mathbf{v}_{h,i}\|_{0,\partial\hat{T}_i}\end{aligned} \tag{2.41}$$

with the mapping $\mathbf{F}_{\hat{T}_i,\mathrm{ref}}$ from Section 2.3. This leads to

$$\langle \mathbf{n}\cdot\mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} \leq h_i^{\frac{1}{2}}\left(\|\mathbf{v}_{h,i}\|_{0,\hat{T}_i}^2 + \|\operatorname{div}\mathbf{v}_{h,i}\|_{0,\hat{T}_i}^2\right)^{\frac{1}{2}}\|q\|_{0,\Gamma_i}\,. \tag{2.42}$$

Summing over all arcs belonging to the boundary leads to

$$\begin{aligned}\langle \mathbf{n}\cdot\mathbf{v}_h, q\rangle_{0,\Gamma} &= \sum_{i=1}^{N}\langle \mathbf{n}\cdot\mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} \\ &\lesssim h^{\frac{1}{2}}\left(\|\mathbf{v}_h\|_{0,\cup_{i=1}^N\hat{T}_i}^2 + \|\operatorname{div}\mathbf{v}_h\|_{0,\cup_{i=1}^N\hat{T}_i}^2\right)^{\frac{1}{2}}\|q\|_{0,\Gamma} \\ &\lesssim h^{\frac{1}{2}}\left(\|\mathbf{v}_h\|_{0,\hat{\Omega}}^2 + \|\operatorname{div}\mathbf{v}_h\|_{0,\hat{\Omega}}^2\right)^{\frac{1}{2}}\|q\|_{0,\Gamma}\end{aligned} \tag{2.43}$$

due to the fact that $\cup_{i=1}^N\hat{T}_i \subset \cup_{i=1}^{\bar{N}}\hat{T}_i = \hat{\Omega}$. Finally, one obtains

$$\langle \mathbf{n}\cdot\mathbf{v}_h, q\rangle_{0,\Gamma} \lesssim h^{\frac{1}{2}}\left(\|\mathbf{v}_h\|_{0,\Omega}^2 + \|\operatorname{div}\mathbf{v}_h\|_{0,\Omega}^2\right)^{\frac{1}{2}}\|q\|_{0,\Gamma}\,. \tag{2.44}$$

using the following equivalence

$$\|\mathbf{v}_h\|_{0,\hat{\Omega}}^2 \eqsim \|\mathbf{v}_h\|_{0,\Omega}^2\,,\ \|\operatorname{div}\mathbf{v}_h\|_{0,\hat{\Omega}}^2 \eqsim \|\operatorname{div}\mathbf{v}_h\|_{0,\Omega}^2 \tag{2.45}$$

that holds due to (2.4) and the fact that $\mathbf{v}_h$ belongs to a finite-dimensional space.

For $q \in H^1(\Gamma)$, starting from (2.39) again and integration by parts leads to

$$\begin{aligned}\langle \mathbf{n}\cdot\mathbf{v}_{h,i}, q\rangle_{0,\Gamma_{\mathbf{D}_i}} &= \int_0^{h_i}\left(-\eta'(\xi)(\alpha+\xi\,\beta)\,q_{\mathbf{D}}(\xi,\eta(\xi)) + \eta(\xi)\,\beta\,q_{\mathbf{D}}(\xi,\eta(\xi))\right)\,d\xi \\ &= \int_0^{h_i}\left(\eta(\xi)(\alpha+\xi\,\beta)\,\frac{d}{d\xi}q_{\mathbf{D}}(\xi,\eta(\xi)) + 2\eta(\xi)\,\beta\,q_{\mathbf{D}}(\xi,\eta(\xi))\right)\,d\xi\end{aligned}$$

due to $\eta(0) = \eta(h_i) = 0$. With the Cauchy-Schwarz inequality, it holds

$$\begin{aligned}\langle \mathbf{n} \cdot \mathbf{v}_{h,i}, q\rangle_{0,\Gamma_{\mathbf{D}_i}} \leq &\left(\int_0^{h_i} \eta(\xi)^2(\alpha + \xi\,\beta)^2 d\xi\right)^{\frac{1}{2}} \left(\int_0^{h_i} \left(\frac{d}{d\xi} q_{\mathbf{D}}(\xi, \eta(\xi))\right)^2 d\xi\right)^{\frac{1}{2}} \\ &+ 2\left(\int_0^{h_i} \eta(\xi)^2 \beta^2 d\xi\right)^{\frac{1}{2}} \left(\int_0^{h_i} q_{\mathbf{D}}(\xi, \eta(\xi))^2 d\xi\right)^{\frac{1}{2}},\end{aligned}$$

which implies

$$\begin{aligned}\langle \mathbf{n} \cdot \mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} &\lesssim h_i^2 \left(\|\mathbf{w}_{h,i}\|_{0,E_i} |q|_{1,\Gamma_i} + \|\operatorname{div} \mathbf{w}_{h,i}\|_{0,E_i} \|q\|_{0,\Gamma_i}\right) \\ &\lesssim h_i^2 \left(\|\mathbf{v}_{h,i}\|^2_{0,\hat{\Gamma}_i} + \|\operatorname{div} \mathbf{v}_{h,i}\|^2_{0,\hat{\Gamma}_i}\right)^{\frac{1}{2}} \left(|q|^2_{1,\Gamma_i} + \|q\|^2_{0,\Gamma_i}\right)^{\frac{1}{2}} \\ &\leq h_i^2 \left(\|\mathbf{v}_{h,i}\|^2_{0,\partial\hat{T}_i} + \|\operatorname{div} \mathbf{v}_{h,i}\|^2_{0,\partial\hat{T}_i}\right)^{\frac{1}{2}} \|q\|_{1,\Gamma_i} \\ &\lesssim h_i^{\frac{3}{2}} \left(\|\mathbf{v}_{h,i}\|^2_{0,\hat{T}_i} + \|\operatorname{div} \mathbf{v}_{h,i}\|^2_{0,\hat{T}_i}\right)^{\frac{1}{2}} \|q\|_{1,\Gamma_i}.\end{aligned} \tag{2.46}$$

using the same scaling arguments as above for $q \in L^2(\Gamma)$. Again, summing over all boundary arcs leads to

$$\begin{aligned}\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma} &\lesssim h^{\frac{3}{2}} \left(\|\mathbf{v}_h\|^2_{0,\hat{\Omega}} + \|\operatorname{div} \mathbf{v}_h\|^2_{0,\hat{\Omega}}\right)^{\frac{1}{2}} \|q\|_{1,\Gamma} \\ &\lesssim h^{\frac{3}{2}} \left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\operatorname{div} \mathbf{v}_h\|^2_{0,\Omega}\right)^{\frac{1}{2}} \|q\|_{1,\Gamma}.\end{aligned} \tag{2.47}$$

The last step is to use the fact that $H^{\frac{1}{2}}(\Gamma)$ is an interpolation space of type $\frac{1}{2}$ for $L^2(\Gamma)$ and $H^1(\Gamma)$ as mentioned in Theorem 1.3. Applying it with $T : q \mapsto \langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma}$, $m = 1$, $C_0 = h^{\frac{1}{2}}$ and $C_1 = h^{\frac{3}{2}}$ leads to $C_{\frac{1}{2}} = h$ and thus to

$$|\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma}| \lesssim h \left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\operatorname{div} \mathbf{v}_h\|^2_{0,\Omega}\right)^{\frac{1}{2}} \|q\|_{\frac{1}{2},\Gamma}. \tag{2.48}$$

□

From equation (2.44) of this proof, the following lemma is extracted, as it is needed in Section 2.5.

Lemma 2.3. *Let Ω have the properties stated in Section 2.1. Then,*

$$|\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma}| \lesssim h^{\frac{1}{2}} \left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\operatorname{div} \mathbf{v}_h\|^2_{0,\Omega}\right)^{\frac{1}{2}} \|q\|_{0,\Gamma} \tag{2.49}$$

holds for all $\mathbf{v}_h \in RT_{0,\hat{\Gamma}}(\Omega)$ and $q \in L^2(\Gamma)$.

2.5 Approximation of the Normalizing Constraint $(p_h, 1)_{0,\Omega} = 0$

As mentioned in Section 2.2, the implementable space $\acute{\mathcal{P}}_1(\Omega)$ can be replaced by $\dot{\mathcal{P}}_1(\Omega)$ in order to simplify the analysis. This is the statement of the following lemma (see [7]):

Lemma 2.4. *Let $\acute{p}_h \in \acute{\mathcal{P}}_1(\Omega)$ and $p_h \in \dot{\mathcal{P}}_1(\Omega)$ be two solutions of the problem (1.36) where the space for p_h is replaced with the corresponding one. Then it holds $p_h = \acute{p}_h - C$ with the constant*

$$C = \frac{(\acute{p}_h, 1)_{0,\Omega}}{(1,1)_{0,\Omega}} \lesssim h. \tag{2.50}$$

Proof: First note that

$$\begin{aligned}(\acute{p}_h,1)_{0,\Omega} &= (\acute{p}_h,1)_{0,\Omega} - (\acute{p}_h,1)_{0,\hat{\Omega}} = (\acute{p}_h,1)_{0,\Omega\setminus\hat{\Omega}} - (\acute{p}_h,1)_{0,\hat{\Omega}\setminus\Omega} \\ &= \sum_{i=1}^{N} \left((\acute{p}_h,1)_{0,\widetilde{T}_i\setminus\hat{T}_i} - (\acute{p}_h,1)_{0,\hat{T}_i\setminus\widetilde{T}_i} \right) ,\end{aligned} \tag{2.51}$$

Now let $\boldsymbol{\chi} : \hat{T}_i \cup \widetilde{T}_i \to \mathbb{R}^2$ be an affine function such that

$$\operatorname{div} \boldsymbol{\chi} = 1 \tag{2.52a}$$

$$\mathbf{n} \cdot \boldsymbol{\chi} = 0 \text{ on } \hat{\Gamma}_i \tag{2.52b}$$

$$(\mathbf{n} \times \boldsymbol{\chi}, 1)_{0,\hat{\Gamma}_i} = 0 \tag{2.52c}$$

Then, using the divergence theorem, it holds

$$\begin{aligned}\left|(\acute{p}_h,1)_{0,\widetilde{T}_i\setminus\hat{T}_i}\right| &= \left|(\acute{p}_h, \operatorname{div}\boldsymbol{\chi})_{0,\widetilde{T}_i\setminus\hat{T}_i}\right| \\ &\le |\langle \acute{p}_h, \mathbf{n}\cdot\boldsymbol{\chi}\rangle_{0,\Gamma_i}| + \left|(\nabla\acute{p}_h, \boldsymbol{\chi})_{0,\widetilde{T}_i\setminus\hat{T}_i}\right| .\end{aligned} \tag{2.53}$$

Due to Lemma 2.3, it holds

$$\begin{aligned}|\langle \acute{p}_h, \mathbf{n}\cdot\boldsymbol{\chi}\rangle_{0,\Gamma_i}| &\lesssim h_i^{1/2} \left(\|\boldsymbol{\chi}\|_{0,\widetilde{T}_i}^2 + \|\operatorname{div}\boldsymbol{\chi}\|_{0,\widetilde{T}_i}^2 \right)^{1/2} \|\acute{p}_h\|_{0,\Gamma_i} \\ &= h_i^{1/2} \left(\|\boldsymbol{\chi}\|_{0,\widetilde{T}_i}^2 + h_i^2 \right)^{1/2} \|\acute{p}_h\|_{0,\Gamma_i} \\ &\lesssim h_i^{1/2} \left(\|\boldsymbol{\chi}\|_{0,\widetilde{T}_i} + h_i \right) \|\acute{p}_h\|_{1,\widetilde{T}_i} ,\end{aligned}$$

where the trace inequality from Theorem 1.2 is used. Combining this with the following scaling argument

$$\|\boldsymbol{\chi}\|_{0,\widetilde{T}_i}^2 \lesssim h_i^2 \|\boldsymbol{\chi}\circ F_{\hat{T}_i,\mathrm{ref}}\|_{0,F_{\hat{T}_i,\mathrm{ref}}^{-1}(\widetilde{T}_i)}^2 \lesssim h_i^2 \tag{2.54}$$

leads to

$$\begin{aligned}\left|(\acute{p}_h,1)_{0,\widetilde{T}_i\setminus\hat{T}_i}\right| &\lesssim h_i^{1/2} \left(h_i^2 + h_i\right) \|\acute{p}_h\|_{1,\widetilde{T}_i} + \|\nabla\acute{p}_h\|_{0,\widetilde{T}_i\setminus\hat{T}_i} \|\boldsymbol{\chi}\|_{0,\widetilde{T}_i\setminus\hat{T}_i} \\ &\lesssim h_i^{3/2} \|\acute{p}_h\|_{1,\widetilde{T}_i} + \|\nabla\acute{p}_h\|_{0,\widetilde{T}_i} \|\boldsymbol{\chi}\|_{0,\widetilde{T}_i} \\ &\lesssim h_i^{3/2} \|\acute{p}_h\|_{1,\widetilde{T}_i} + h_i^2 \|\acute{p}_h\|_{1,\widetilde{T}_i}\end{aligned}$$

and thus to

$$\left|(\acute{p}_h,1)_{0,\widetilde{T}_i\setminus\hat{T}_i}\right| \lesssim h_i^{3/2} \|\acute{p}_h\|_{1,\widetilde{T}_i}. \tag{2.55}$$

The estimate

$$\left|(\acute{p}_h,1)_{0,\hat{T}_i\setminus\widetilde{T}_i}\right| \lesssim h_i^{3/2} \|\acute{p}_h\|_{1,\widetilde{T}_i} \tag{2.56}$$

is obtained in a similar way. Insert in (2.51) to get

$$\begin{aligned} |(\acute{p}_h, 1)_{0,\Omega}| \leq & \sum_{i=1}^{N} \left((\acute{p}_h, 1)_{0,\widetilde{T}_i \setminus \hat{T}_i} - (\acute{p}_h, 1)_{0,\hat{T}_i \setminus \widetilde{T}_i} \right) \\ \lesssim & \sum_{i=1}^{N} h_i^{3/2} \|\acute{p}_h\|_{1,\widetilde{T}_i} \\ \leq & h \left(\sum_{i=1}^{N} h_i \right)^{1/2} \left(\|\acute{p}_h\|_{1,\widetilde{T}}^2 \right)^{1/2} \end{aligned} \tag{2.57}$$

and using the fact that

$$\sum_{i=1}^{N} h_i = |\hat{\Gamma}| \,,$$

leads to the result:

$$|(\acute{p}_h, 1)_{0,\Omega}| \lesssim h \|\acute{p}_h\|_{1,\Omega} \,, \tag{2.58}$$

□

Chapter 3

Least Squares Method on Domains with Curved Boundaries

The aim of this Chapter is to show that the use of the polygonal approximation of the boundary derived in chapter 2 is sufficient to retain the optimal order of convergence for the lowest order elements derived in Section 2.2.

The first step therefore is to construct an implementable approximation of the least squares functional $\mathcal{F}$, and this will be done in the first section, in particular the $L^2(\hat{\Omega})$ orthogonal projection on the piecewise constant functions is used to approximate the right side f.

In the second section, the main theorem of this chapter is then proved. To this end, a lower bound for the least squares functional is derived using the estimate for the normal flux of the finite-element funtion on the approximated boundary from Theorem 2.1. Then, a connection between the least squares functional and its approximation is established and finally an upper bound for the approximated functional is stated, using the interpolation properties of the finite-element interpolation operators derived in Section 1.5.

This theoretical result is then illustrated by a numerical example in the third section.

3.1 Approximation of the Least Squares Functional

After the construction of the curved triangulation and the definition of the corresponding finite elements, the next step to apply the least squares method on domains with curved boundaries is to define an implementable approximated functional $\mathcal{F}_h$. Therefore, the integration domain Ω is replaced by $\hat{\Omega}$. Further, the function $f \in L^2(\Omega)$ has to be replaced by an implementable one, such that

$$\mathcal{F}_h(\mathbf{u}_h, p_h) = \|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,\hat{\Omega}} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\hat{\Omega}} \,. \tag{3.1}$$

has to be minimized with respect to $\mathbf{u}_h \in RT_{0,\hat{\Gamma}}(\Omega)$ and $p_h \in \dot{\mathcal{P}}_1(\Omega)$.

The following lemma (published in [7]) states that choosing f_h as the $L^2(\hat{\Omega})$-orthogonal projection onto piecewise constant functions leads to a sufficiently good approximation to f on Ω.

Lemma 3.1. *Let f belong to $W^1_\infty(\hat{\Omega} \cup \widetilde{\Omega})$ and $\mathcal{P}_h : L^2(\hat{\Omega}) \to Z_h$ be the $L^2(\hat{\Omega})$-orthogonal projection onto the space Z_h of piecewise constants on the triangulation $\widehat{\mathcal{T}}_h$ of $\hat{\Omega}$. Then,*

$$\|f - \mathcal{P}_h f\|_{0,\Omega} \lesssim h \|f\|_{W^1_\infty(\Omega \cup \hat{\Omega})} \tag{3.2}$$

holds for any $f \in W^1_\infty(\Omega \cup \hat{\Omega})$

Proof: For each $\hat{T}_i \in \hat{\mathcal{T}}_h$, due to the orthogonality with respect to $L^2(\hat{T}_i)$ it holds

$$(\mathcal{P}_h f, 1)_{0,\hat{T}_i} = (f,1)_{0,\hat{T}_i} \tag{3.3}$$

and for each $\widetilde{T} \in \widetilde{\mathcal{T}}_h$, due to the orthogonality with respect to $L^2(\widetilde{T}_i)$ it holds

$$(\tilde{\mathcal{P}}_h f, 1)_{0,\widetilde{T}_i} = (f,1)_{0,\widetilde{T}_i} \tag{3.4}$$

where $\tilde{\mathcal{P}}_h : L^2(\Omega) \to Z_h$ denotes the $L^2(\Omega)$-orthogonal projection. Due to the fact that $\mathcal{P}_h f\big|_{\hat{T}_i}$ and $\widetilde{\mathcal{P}}_h f\big|_{\widetilde{T}_i}$ are constants one obtains

$$\mathcal{P}_h f\big|_{\widetilde{T}_i} = \frac{(f,1)_{0,\hat{T}_i}}{(1,1)_{0,\hat{T}_i}} \text{ and } \widetilde{\mathcal{P}}_h f\big|_{\widetilde{T}_i} = \frac{(f,1)_{0,\widetilde{T}_i}}{(1,1)_{0,\widetilde{T}_i}} \,. \tag{3.5}$$

Further, note that

$$\begin{aligned}
\|f - \tilde{\mathcal{P}}_h f\|^2_{0,\widetilde{T}_i} &= \|f\|_{0,\widetilde{T}_i} - (f, \tilde{\mathcal{P}}_h f)_{0,\widetilde{T}_i} - (f, \tilde{\mathcal{P}}_h f)_{0,\widetilde{T}_i} + (\tilde{\mathcal{P}}_h f, \tilde{\mathcal{P}}_h f)_{0,\widetilde{T}_i} \\
&= \|f\|^2_{0,\widetilde{T}_i} - (f, \tilde{\mathcal{P}}_h f)_{0,\widetilde{T}_i} \\
&= \|f\|^2_{0,\widetilde{T}_i} - \frac{(f,1)^2_{0,\widetilde{T}_i}}{(1,1)_{0,\widetilde{T}_i}}
\end{aligned}$$

such that $\|f - \tilde{\mathcal{P}}_h f\|^2_{0,\widetilde{T}_i}$ and $\|f - \mathcal{P}_h f\|^2_{0,\widetilde{T}_i}$ can be connected:

$$\begin{aligned}
\|f - \mathcal{P}_h f\|^2_{0,\widetilde{T}_i} &= \|f\|^2_{0,\widetilde{T}_i} - 2\mathcal{P}_h f\big|_{\widetilde{T}_i}(f,1)_{0,\widetilde{T}_i} + (\mathcal{P}_h f, \mathcal{P}_h f)_{0,\widetilde{T}_i} \\
&= \|f - \widetilde{\mathcal{P}}_h f\|^2_{0,\widetilde{T}_i} + \frac{(f,1)^2_{0,\widetilde{T}_i}}{(1,1)_{0,\widetilde{T}_i}} - 2\mathcal{P}_h f\big|_{\widetilde{T}_i}(f,1)_{0,\widetilde{T}_i} + \mathcal{P}_h f\big|^2_{\widetilde{T}_i}(1,1)_{0,\widetilde{T}_i} \\
&= \|f - \widetilde{\mathcal{P}}_h f\|^2_{0,\widetilde{T}_i} + (1,1)_{0,\widetilde{T}_i}\left(\frac{(f,1)^2_{0,\widetilde{T}_i}}{(1,1)^2_{0,\widetilde{T}_i}} - 2\mathcal{P}_h f\big|_{\widetilde{T}_i}\frac{(f,1)_{0,\widetilde{T}_i}}{(1,1)_{0,\widetilde{T}_i}} + \mathcal{P}_h f\big|^2_{\widetilde{T}_i}\right) \\
&= \|f - \widetilde{\mathcal{P}}_h f\|^2_{0,\widetilde{T}_i} + (1,1)_{0,\widetilde{T}_i}\left(\widetilde{\mathcal{P}}_h f\big|_{\widetilde{T}_i} - \mathcal{P}_h f\big|_{\widetilde{T}_i}\right)^2
\end{aligned} \tag{3.6}$$

With the best approximation property

$$\|f - \widetilde{\mathcal{P}}_h f\|^2_{0,\widetilde{T}_i} \lesssim h_i^2 \|f\|^2_{1,\widetilde{T}_i} \tag{3.7}$$

of $\widetilde{\mathcal{P}}_h f$ on $\widetilde{T}_i$, this leads to

$$\|f - \mathcal{P}_h f\|^2_{0,\widetilde{T}_i} \lesssim h_i^2 \|f\|^2_{1,\widetilde{T}_i} + |\widetilde{T}_i|\left(\widetilde{\mathcal{P}}_h f\big|_{\widetilde{T}_i} - \mathcal{P}_h f\big|_{\widetilde{T}_i}\right)^2 \tag{3.8}$$

It remains to estimate $\left|\widetilde{\mathcal{P}}_h f\big|_{\widetilde{T}_i} - \mathcal{P}_h f\big|_{\widetilde{T}_i}\right|$. To this end, first use the transformation $\mathbf{D}_i$ of the proof of Theorem 2.1 and the function η to note that

$$\left|(1,1)_{0,\hat{T}_i} - (1,1)_{0,\widetilde{T}_i}\right| = \left|\left|\hat{T}_i\right| - \left|\widetilde{T}_i\right|\right| = \left|\int_0^{h_i} \eta(\xi)\ \mathrm{d}\xi\right| \lesssim \int_0^{h_i} |\eta(\xi)|\ \mathrm{d}\xi \lesssim h_i^3 \tag{3.9}$$

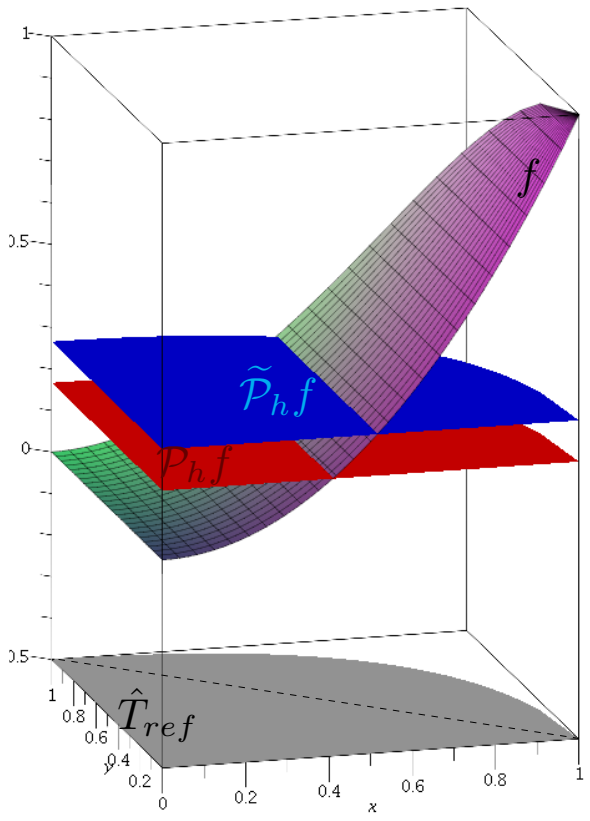


Figure 3.1: Example for the difference between $\widetilde{\mathcal{P}}_h f$ and $\mathcal{P}_h f$

holds, and thus

$$\begin{aligned}
\left|\widetilde{\mathcal{P}}_h f\right|_{\widetilde{T}_i} - \mathcal{P}_h f\Big|_{\widetilde{T}_i}\Big| &= \left|\frac{(f,1)_{0,\widetilde{T}_i}}{(1,1)_{0,\widetilde{T}_i}} - \frac{(f,1)_{0,\hat{T}_i}}{(1,1)_{0,\hat{T}_i}}\right| \\
&= \left|\frac{(f,1)_{0,\widetilde{T}_i}((1,1)_{0,\hat{T}_i} - (1,1)_{0,\widetilde{T}_i})}{(1,1)_{0,\widetilde{T}_i}(1,1)_{0,\hat{T}_i}} + \frac{(f,1)_{0,\widetilde{T}_i} - (f,1)_{0,\hat{T}_i}}{(1,1)_{0,\hat{T}_i}}\right| \\
&\lesssim h_i^{-4}\left|(f,1)_{0,\widetilde{T}_i}((1,1)_{0,\hat{T}_i} - (1,1)_{0,\widetilde{T}_i})\right| + h_i^{-2}\left|(f,1)_{0,\widetilde{T}_i} - (f,1)_{0,\hat{T}_i}\right| \\
&\lesssim h_i^{-1}\left|(f,1)_{0,\widetilde{T}_i}\right| + h_i^{-2}\left|(f,1)_{0,\widetilde{T}_i} - (f,1)_{0,\hat{T}_i}\right|.
\end{aligned} \tag{3.10}$$

Using lemma 2.2, this turns to

$$\left|\widetilde{\mathcal{P}}_h f\right|_{\widetilde{T}_i} - \mathcal{P}_h f\Big|_{\widetilde{T}_i}\Big| \lesssim h_i^{-1}\left|(f,1)_{0,\widetilde{T}_i}\right| + h_i\|f\|_{W^1_\infty(\hat{T}_i\cup\widetilde{T}_i)} + \|f\|_{0,\widetilde{T}_i} \tag{3.11}$$

and then to

$$\left|\widetilde{\mathcal{P}}_h f\right|_{\widetilde{T}_i} - \mathcal{P}_h f\Big|_{\widetilde{T}_i}\Big| \lesssim h_i\|f\|_{W^1_\infty(\hat{T}_i\cup\widetilde{T}_i)} + \|f\|_{0,\widetilde{T}_i}\,, \tag{3.12}$$

using the Cauchy-Schwarz inequality

$$\left|(f,1)_{0,\widetilde{T}_i}\right|^2 \le (f,f)_{0,\widetilde{T}_i}(1,1)_{0,\widetilde{T}_i} \lesssim h^2\|f\|^2_{0,\widetilde{T}_i}\,. \tag{3.13}$$

Now, inserting (3.12) into (3.8) leads to

$$\begin{aligned}\|f-\mathcal{P}_h f\|_{0,\widetilde{T}_i}^2 \lesssim & h_i^2\|f\|_{1,\widetilde{T}_i}^2 + h^2\left(h_i\|f\|_{W_\infty^1(\hat{T}_i\cup\widetilde{T}_i)} + \|f\|_{0,\widetilde{T}_i}\right)^2 \\ \lesssim & h_i^2\left(\|f\|_{1,\widetilde{T}_i}^2 + h_i^2\|f\|_{W_\infty^1(\hat{T}_i\cup\widetilde{T}_i)}^2\right) \\ \lesssim & h_i^2\|f\|_{W_\infty^1(\hat{T}_i\cup\widetilde{T}_i)}^2\end{aligned} \tag{3.14}$$

where the imbedding inequality that follows from (1.20a) is used. Now, summing over all triangles and using the best approximation property on the interior triangles leads to

$$\begin{aligned}\|f-\mathcal{P}_h f\|_{0,\Omega} &= \sum_{i=1}^{\bar{N}} \|f-\mathcal{P}_h f\|_{0,\widetilde{T}_i} \\ &= \sum_{i=N+1}^{\bar{N}} \|f-\mathcal{P}_h f\|_{0,\hat{T}_i} + \sum_{i=1}^{N} \|f-\mathcal{P}_h f\|_{0,\widetilde{T}_i} \\ &\lesssim \sum_{i=N+1}^{\bar{N}} h_i\|f\|_{1,\hat{T}_i} + \sum_{i=1}^{N} h_i\|f\|_{W_\infty^1(\hat{T}_i\cup\widetilde{T}_i)}\end{aligned} \tag{3.15}$$

such that using (1.20a) again leads to the result:

$$\begin{aligned}\|f-\mathcal{P}_h f\|_{0,\Omega} &\lesssim \sum_{i=N+1}^{\bar{N}} h_i\|f\|_{W_\infty^1(\hat{T}_i)} + \sum_{i=1}^{N} h_i\|f\|_{W_\infty^1(\hat{T}_i\cup\widetilde{T}_i)} \\ &\lesssim h\|f\|_{W_\infty^1(\Omega\cup\hat{\Omega})}.\end{aligned} \tag{3.16}$$

□

3.2 Error Analysis of the Least Squares FEM on Domains with Curved Boundaries

The aim of this section is to derive a theorem similar to Theorem 1.9 for the domain Ω with curved boundaries considered from Chapter 2 on. A key point therefore is to derive a lower bound of the least squares functional, as the bound derived in Section 1.5 does not hold due to the fact that the normal flux does not vanish on Γ. However, an upper bound can be given using the estimate from Theorem 2.1, as the following lemma (see [7]) states:

Lemma 3.2. *Let $(\mathbf{u},p)\in H_\Gamma(\mathrm{div},\Omega)\times\dot{H}^1(\Omega)$ denote the exact solution of the system (1.33). Then, it holds*

$$\mathcal{F}(\mathbf{v}_h,q_h)+h^2\left(\|\mathbf{u}\|_{0,\Omega}^2+\|\mathrm{div}\,\mathbf{u}\|_{0,\Omega}^2\right)\gtrsim\|\mathbf{v}_h-\mathbf{u}\|_{\mathrm{div},\Omega}^2+\|\nabla(q_h-p)\|_{0,\Omega}^2 \tag{3.17}$$

for all $(\mathbf{v}_h,q_h)\in RT_{0,\hat{\Gamma}}(\Omega)\times\dot{\mathcal{P}}_1(\Omega)$.

Proof: First note that due to the fact that the exact solution solve (1.33), it can be inserted in $\mathcal{F}(\mathbf{v}_h,q_h)$:

$$\begin{aligned}\mathcal{F}(\mathbf{v}_h,q_h) &= \|\mathrm{div}\,\mathbf{v}_h-f\|_{0,\Omega}^2+\|\mathbf{v}_h+\nabla q_h\|_{0,\Omega}^2 \\ &= \|\mathrm{div}\,\mathbf{v}_h-f-\mathrm{div}\,\mathbf{u}+f\|_{0,\Omega}^2+\|\mathbf{v}_h+\nabla q_h-\mathbf{u}-\nabla p\|_{0,\Omega}^2 \\ &= \|\mathrm{div}(\mathbf{v}_h-\mathbf{u})\|_{0,\Omega}^2+\|\mathbf{v}_h-\mathbf{u}+\nabla(q_h-p)\|_{0,\Omega}^2\end{aligned}$$

As in Section 1.5, the idea to get the lower bound is to expand the term $\|\mathbf{v}_h - \mathbf{u} + \nabla(q_h - p)\|^2_{0,\Omega}$ and to distribute the mixed terms that then appears on the other terms, and in particular to weight it against $\|\mathrm{div}(\mathbf{v}_h - \mathbf{u})\|^2_{0,\Omega}$:

$$\begin{aligned}\mathcal{F}(\mathbf{v}_h, q_h) = \|\mathrm{div}(\mathbf{v}_h - \mathbf{u})\|^2_{0,\Omega} + \|\mathbf{v}_h - \mathbf{u}\|^2_{0,\Omega} + \|\nabla(q_h - p)\|^2_{0,\Omega} \\ - 2(1-\alpha)(\mathbf{v}_h - \mathbf{u}, \nabla(q_h - p))_{0,\Omega} \\ + 2\alpha\left(\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma} - (\mathrm{div}(\mathbf{v}_h - \mathbf{u}), q_h - p)_{0,\Omega}\right)\end{aligned}$$

with any $\alpha \in [0,1]$. Using

$$\begin{aligned}-|2(1-\alpha)(\mathbf{v}_h - \mathbf{u}, \nabla(q_h - p))_{0,\Omega}| &\geq (\alpha - 1)\left(\|\mathbf{v}_h - \mathbf{u}\|^2_{0,\Omega} + \|\nabla(q_h - p)\|^2_{0,\Omega}\right) \\ \text{and} \quad -|2\alpha(\mathrm{div}(\mathbf{v}_h - \mathbf{u}), q_h - p)_{0,\Omega}| &\geq \alpha\left(-\frac{1}{2\alpha}\|\mathrm{div}(\mathbf{v}_h - \mathbf{u})\|^2_{0,\Omega} - 2\alpha\|q_h - p\|^2_{0,\Omega}\right)\end{aligned}$$

leads to

$$\begin{aligned}\mathcal{F}(\mathbf{v}_h, q_h) \geq \frac{1}{2}\|\mathrm{div}(\mathbf{v}_h - \mathbf{u})\|^2_{0,\Omega} + \alpha\|\mathbf{v}_h - \mathbf{u}\|^2_{0,\Omega} + \alpha\|\nabla(q_h - p)\|^2_{0,\Omega} \\ - 2\alpha^2\|q_h - p\|^2_{0,\Omega} - 2\alpha\,|\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma}| \ .\end{aligned} \tag{3.18}$$

Using the Poincaré-inequality (1.25), this becomes

$$\begin{aligned}\mathcal{F}(\mathbf{v}_h, q_h) \geq \frac{1}{2}\|\mathrm{div}(\mathbf{v}_h - \mathbf{u})\|^2_{0,\Omega} + \alpha\|\mathbf{v}_h - \mathbf{u}\|^2_{0,\Omega} \\ + \alpha(1 - 2\alpha C_\Omega)\|\nabla(q_h - p)\|^2_{0,\Omega} - 2\alpha\,|\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma}| \ .\end{aligned} \tag{3.19}$$

The boundary term $|\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma}|$ can then be bounded using Theorem 2.1:

$$2\,|\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma}| \leq 2C_S h\left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right)^{1/2}\|q_h - p\|_{1/2,\Gamma} \ .$$

In order to weight $\|q_h - p\|_{1/2,\Gamma}$ against the term $\alpha(1 - 2\alpha C_\Omega)\|\nabla(q_h - p)\|_{0,\Gamma}$ from the least squares functional, the Cauchy Schwarz inequality and the combined inequality (1.26) have to be used. However, before using the Cauchy Schwarz inequality, the terms have to be weighted correctly:

$$\begin{aligned}2\,|\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma}| &\leq \left(8C_T^2C_\Omega^2C_S^2h^2\|\mathbf{v}_h\|^2_{0,\Omega} + \|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right)^{\frac{1}{2}}\left(\frac{1}{2C_T^2C_\Omega^2}\|q_h - p\|^2_{1/2,\Gamma}\right)^{\frac{1}{2}} \\ &\leq Ch^2\left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right) + \frac{1}{2C_T^2C_\Omega^2}\|q_h - p\|^2_{1/2,\Gamma}\end{aligned}$$

with C=$8C_T^2C_\Omega^2C_S^2$, such that combining the trace inequality and the Poincaré Friedrichs inequality (see (1.26)) leads to

$$2\,|\langle \mathbf{n}\cdot\mathbf{v}_h, q_h - p\rangle_{0,\Gamma}| \leq Ch^2\left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right) + \frac{1}{2}\|\nabla(q_h - p)\|^2_{0,\Omega}\,.$$

Inserting this in (3.19) leads to

$$\begin{aligned}\mathcal{F}(\mathbf{v}_h, q_h) \geq -\alpha Ch^2\left(\|\mathbf{v}_h\|^2_{0,\Omega} + \|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right) + \frac{1}{2}\|\mathrm{div}(\mathbf{v}_h - \mathbf{u})\|^2_{0,\Omega} + \alpha\|\mathbf{v}_h - \mathbf{u}\|^2_{0,\Omega} \\ + \alpha\left(\frac{1}{2} - 2\alpha C_\Omega\right)\|\nabla(q_h - p)\|^2_{0,\Omega}\end{aligned} \tag{3.20}$$

Choosing $\alpha = \min(\frac{1}{2}, \frac{1}{8C_\Omega})$ implies

$$\mathcal{F}(\mathbf{v}_h, q_h) \geq -Ch^2 \left(\|\mathbf{v}_h\|_{0,\Omega}^2 + \|\operatorname{div} \mathbf{v}_h\|_{0,\Omega}^2 \right) + \alpha \|\mathbf{v}_h - \mathbf{u}\|_{\operatorname{div},\Omega}^2 + \frac{\alpha}{4} \|\nabla(q_h - p)\|_{0,\Omega}^2 \tag{3.21}$$

and thus

$$\mathcal{F}(\mathbf{v}_h, q_h) + Ch^2 \left(\|\mathbf{v}_h\|_{0,\Omega}^2 + \|\operatorname{div} \mathbf{v}_h\|_{0,\Omega}^2 \right) \geq \frac{\alpha}{4} \left(\|\mathbf{v}_h - \mathbf{u}\|_{\operatorname{div},\Omega}^2 + \|\nabla(q_h - p)\|_{0,\Omega}^2 \right) \tag{3.22}$$

Consequently, using $\|\mathbf{u}_h\|_{0,\Omega}^2 \leq 2\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}^2 + 2\|\mathbf{u}\|_{0,\Omega}^2$,

$$\mathcal{F}(\mathbf{v}_h, q_h) + h^2 \left(\|\mathbf{u}\|_{0,\Omega}^2 + \|\operatorname{div} \mathbf{u}\|_{0,\Omega}^2 \right) \gtrsim \|\mathbf{v}_h - \mathbf{u}\|_{\operatorname{div},\Omega}^2 + \|\nabla(q_h - p)\|_{0,\Omega}^2 \tag{3.23}$$

holds for $h \leq h_0 = \frac{\alpha - 4}{8C}$. Inequality (3.23) also holds for $h \geq h_0$, since the left-hand side in (3.23) is bounded from below by $\mathcal{F}(\mathbf{v}_h, q_h) + h_0^2 \left(\|\mathbf{u}\|_{0,\Omega}^2 + \|\operatorname{div} \mathbf{u}\|_{0,\Omega}^2 \right)$. Thus, (3.23) holds for all h. □

On the other hand, an upper bound has to be found for the least squares functional

$$\mathcal{F}_h(\mathbf{u}_h, p_h) = \|\operatorname{div} \mathbf{u}_h - f_h\|_{0,\hat{\Omega}}^2 + \|\mathbf{u}_h + \nabla p_h\|_{0,\hat{\Omega}}^2 . \tag{3.24}$$

where f_h is the $L^2(\hat{\Omega})$-orthogonal projection of f onto the space of piecewise constant functions. This is the statement of the following Lemma (see [7]).

Lemma 3.3. *Let $(\mathbf{u}, p) \in H_\Gamma(\operatorname{div}, \Omega) \times \dot{H}^1(\Omega)$ denote the exact solution of the system (1.33) and assume that it satisfies $p \in H^2(\Omega)$. Further, let $(\mathbf{u}_h, p_h)$ be the minimizer of $\mathcal{F}_h(\mathbf{v}_h, q_h)$ under all $(\mathbf{v}_h, p_h) \in RT_{0,\hat{\Gamma}}(\Omega) \times \dot{\mathcal{P}}_1(\Omega)$. Then, it holds*

$$\mathcal{F}_h(\mathbf{u}_h, p_h) \lesssim h^2 \left(\|\mathbf{u}\|_{1,\Omega}^2 + \|p\|_{2,\Omega}^2 \right) . \tag{3.25}$$

Proof: The idea is to use the fact that the solution $(\mathbf{u}_h, p_h)$ minimize the approximated least squares functional and thus to bound the least squares functional with the interpolation of the exact solution. Therefore, first map the exact solution on $\hat{\Omega}$, using the mapping $\hat{\mathbf{\Psi}}_h : \Omega \to \hat{\Omega}$ from Lemma 2.1:

$$\begin{aligned} \tilde{\mathbf{u}} &= \mathcal{G}_{Piola}(\mathbf{u}) = \left(\frac{1}{\det J_{\hat{\mathbf{\Psi}}_h}} J_{\hat{\mathbf{\Psi}}_h} \mathbf{u} \right) \circ \hat{\mathbf{\Phi}}_h , \\ \tilde{p} &= p \circ \hat{\mathbf{\Phi}}_h \end{aligned} \tag{3.26}$$

Note that then, $\tilde{\mathbf{u}} \in H_{\hat{\Gamma}}(\operatorname{div}, \hat{\Omega})$ and $\tilde{p} \in H^1(\hat{\Omega})$. Using the fact that $(\mathbf{u}_h, p_h)$ minimizes $\mathcal{F}_h(\mathbf{v}_h, q_h)$ in the finite-element space $RT_{0,\hat{\Gamma}}(\Omega) \times \dot{\mathcal{P}}_1(\Omega)$ and the interpolations operator $\hat{\mathcal{R}}_h : H_{\hat{\Gamma}}(\operatorname{div}, \hat{\Omega}) \to RT_{0,\hat{\Gamma}}(\hat{\Omega})$ and $\mathcal{I}_h : \dot{H}^1(\hat{\Omega}) \to \dot{\mathcal{P}}_1(\hat{\Omega})$ from Section 1.5 leads to

$$\mathcal{F}_h(\mathbf{u}_h, p_h) \leq \mathcal{F}_h(\hat{\mathcal{R}}_h \tilde{\mathbf{u}}, \mathcal{I}_h \tilde{p}) = \|\operatorname{div}(\hat{\mathcal{R}}_h \tilde{\mathbf{u}}) - f_h\|_{0,\hat{\Omega}}^2 + \|\hat{\mathcal{R}}_h \tilde{\mathbf{u}} + \nabla(\mathcal{I}_h \tilde{p})\|_{0,\hat{\Omega}}^2 \tag{3.27}$$

The first term vanishes due to the definition of f_h as the $L^2(\hat{\Omega})$ projection of f and due to the fact that the Raviart-Thomas interpolation of the divergence coincides with its L^2 projection (see Proposition 2.5.2 in [11]):

$$\|\operatorname{div}(\hat{\mathcal{R}}_h \tilde{\mathbf{u}}) - f_h\|_{0,\hat{\Omega}}^2 = \|\mathcal{P}_h(\operatorname{div} \mathbf{u}) - \mathcal{P}_h f\|_{0,\hat{\Omega}}^2 = \|\mathcal{P}_h(\operatorname{div} \mathbf{u} - f)\|_{0,\hat{\Omega}}^2 = 0 \tag{3.28}$$

Thus,

$$\begin{aligned}\mathcal{F}_h(\mathbf{u}_h, p_h) &= \|\hat{\mathcal{R}}_h\tilde{\mathbf{u}} + \nabla(\mathcal{I}_h\tilde{p})\|^2_{0,\hat{\Omega}} \\ &= \|\hat{\mathcal{R}}_h\tilde{\mathbf{u}} - \tilde{\mathbf{u}} + \tilde{\mathbf{u}} + \nabla\tilde{p} - \nabla\tilde{p} + \nabla(\mathcal{I}_h\tilde{p})\|^2_{0,\hat{\Omega}} \\ &= \|\hat{\mathcal{R}}_h\tilde{\mathbf{u}} - \tilde{\mathbf{u}}\|^2_{0,\hat{\Omega}} + \|\tilde{\mathbf{u}} + \nabla\tilde{p}\|^2_{0,\hat{\Omega}} + \|\nabla(\mathcal{I}_h\tilde{p} - \tilde{p})\|^2_{0,\hat{\Omega}} \,.\end{aligned} \tag{3.29}$$

The approximation results for the finite-elements from Section 1.5 leads to

$$\mathcal{F}_h(\mathbf{u}_h, p_h) \lesssim h^2\|\tilde{\mathbf{u}}\|^2_{1,\hat{\Omega}} + h^2\|\tilde{p}\|^2_{2,\hat{\Omega}} + \|\tilde{\mathbf{u}} + \nabla\tilde{p}\|^2_{0,\hat{\Omega}} \tag{3.30}$$

Now, $\|\tilde{\mathbf{u}} + \nabla\tilde{p}\|^2_{0,\hat{\Omega}}$ has to be bounded. This can be done using the mapping $\hat{\boldsymbol{\Psi}}_h$ from Lemma 2.1 that implies

$$\begin{aligned}\|\tilde{\mathbf{u}} + \nabla\tilde{p}\|^2_{0,\hat{\Omega}} &= \int_{\hat{\Omega}} (\tilde{\mathbf{u}} + \nabla\tilde{p})^2 \, \mathbf{d}\hat{\mathbf{x}} \\ &= \int_{\Omega} \left(\frac{1}{\det J_{\hat{\boldsymbol{\Psi}}_h}} J_{\hat{\boldsymbol{\Psi}}_h}\mathbf{u} + J^T_{\hat{\boldsymbol{\Phi}}_h} \nabla p \right)^2 \det J_{\hat{\boldsymbol{\Psi}}_h} \, \mathbf{dx} \\ &= \int_{\Omega} \left(\left(\frac{1}{\det J_{\hat{\boldsymbol{\Psi}}_h}} J_{\hat{\boldsymbol{\Psi}}_h} - I \right)\mathbf{u} + \left(J^T_{\hat{\boldsymbol{\Phi}}_h} - I \right)\nabla p \right)^2 \det J_{\hat{\boldsymbol{\Psi}}_h} \, \mathbf{dx}\end{aligned}$$

such that using Inequality (2.22) leads to

$$\begin{aligned}\|\tilde{\mathbf{u}} + \nabla\tilde{p}\|^2_{0,\hat{\Omega}} &\lesssim \left\| \frac{1}{\det J_{\hat{\boldsymbol{\Psi}}_h}} J_{\hat{\boldsymbol{\Psi}}_h} - I \right\|^2_{L_\infty(\Omega)} \|\mathbf{u}\|^2_{0,\Omega} + \|J_{\hat{\boldsymbol{\Phi}}_h} - I\|_{L_\infty(\Omega)} \|\nabla p\|^2_{0,\Omega} \\ &\lesssim h^2 \left(\|\mathbf{u}\|^2_{0,\Omega} + \|\nabla p\|^2_{0,\Omega} \right) .\end{aligned} \tag{3.31}$$

Inserting (3.31) into (3.30) implies

$$\mathcal{F}_h(\mathbf{u}_h, p_h) \lesssim h^2 \left(\|\tilde{\mathbf{u}}\|^2_{1,\hat{\Omega}} + \|\tilde{p}\|^2_{2,\hat{\Omega}} \right) + h^2 \left(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{2,\Omega} \right) . \tag{3.32}$$

Using the Theorems 1.6 and 1.8 leads to the result

$$\mathcal{F}_h(\mathbf{u}_h, p_h) \lesssim h^2 \left(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{2,\Omega} \right) . \tag{3.33}$$

□

The next theorem, which is the main result in [7], shows that the optimal order convergence of the least squares method described in chapter 1 is retained for domains with curved boundaries.

Theorem 3.1. *Let $(\mathbf{u}, p) \in H_\Gamma(\mathrm{div}, \Omega) \times \dot{H}^1(\Omega)$ denote the exact solution of the system (1.33) and assume that it satisfies $p \in H^2(\Omega)$. Further, assume $f \in W^1_\infty(\hat{\Omega} \cup \widetilde{\Omega})$ and let $(\mathbf{u}_h, p_h)$ be the minimizer of $\mathcal{F}_h(\mathbf{v}_h, q_h)$ under all $(\mathbf{v}_h, p_h) \in RT_{0,\hat{\Gamma}}(\Omega) \times \dot{\mathcal{P}}_1(\Omega)$. Then,*

$$\|\mathbf{u} - \mathbf{u}_h\|_{\mathrm{div},\Omega} + \|p - p_h\|_{1,\Omega} \lesssim h \left(\|\mathbf{u}\|_{1,\Omega} + \|f\|_{W^1_\infty(\Omega\cup\hat{\Omega})} + \|p\|_{2,\Omega} \right) \tag{3.34}$$

holds.

Proof: First recall that from the previous lemma it holds

$$\mathcal{F}_h(\mathbf{u}_h, p_h) \lesssim h^2 \left(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{2,\Omega} \right) \tag{3.35}$$

$$\text{and } \mathcal{F}(\mathbf{u}_h, p_h) + h^2 \left(\|\mathbf{u}\|^2_{0,\Omega} + \|\operatorname{div} \mathbf{u}\|^2_{0,\Omega} \right) \gtrsim \|\mathbf{u}_h - \mathbf{u}\|^2_{\operatorname{div},\Omega} + \|\nabla(p_h - p)\|^2_{0,\Omega} \tag{3.36}$$

Clearly, $\mathcal{F}_h(\mathbf{u}_h, p_h)$ and $\mathcal{F}(\mathbf{v}_h, q_h)$ need to be connected. This can be done using the intermediate functional

$$\widetilde{\mathcal{F}}_h(\mathbf{u}_h, p_h) = \|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,\Omega} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\Omega} \,. \tag{3.37}$$

that leads on the one hand to

$$\mathcal{F}(\mathbf{u}_h, p_h) = \|f - f_h\|^2_{0,\Omega} + \widetilde{\mathcal{F}}_h(\mathbf{u}_h, p_h) \lesssim h^2 \|f\|^2_{W^1_\infty(\Omega \cup \hat{\Omega})} + \widetilde{\mathcal{F}}_h(\mathbf{u}_h, p_h) \,, \tag{3.38}$$

using Lemma 3.1, and on the other hand

$$\widetilde{\mathcal{F}}_h(\mathbf{u}_h, p_h) = \mathcal{F}_h(\mathbf{u}_h, p_h) + \|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,\Omega\setminus\hat{\Omega}} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\Omega\setminus\hat{\Omega}} \eqsim \mathcal{F}_h(\mathbf{u}_h, p_h) \tag{3.39}$$

due to (2.4) and the fact that $(\mathbf{u}_h, p_h)$ belongs to a finite dimensional space. An upper bound for the least squares functional is then obtained combining (3.35), (3.38) and (3.39) :

$$\begin{aligned} \mathcal{F}(\mathbf{u}_h, p_h) &\lesssim h^2 \|f\|^2_{W^1_\infty(\Omega \cup \hat{\Omega})} + \mathcal{F}_h(\mathbf{u}_h, p_h) \\ &\lesssim h^2 \|f\|^2_{W^1_\infty(\Omega \cup \hat{\Omega})} + h^2 \left(\|\mathbf{u}\|^2_{1,\Omega} + \|p\|^2_{2,\Omega} \right) \end{aligned} \tag{3.40}$$

The result finally follows from combining (3.36) and (3.40) with the Poincaré inequality (1.25). □

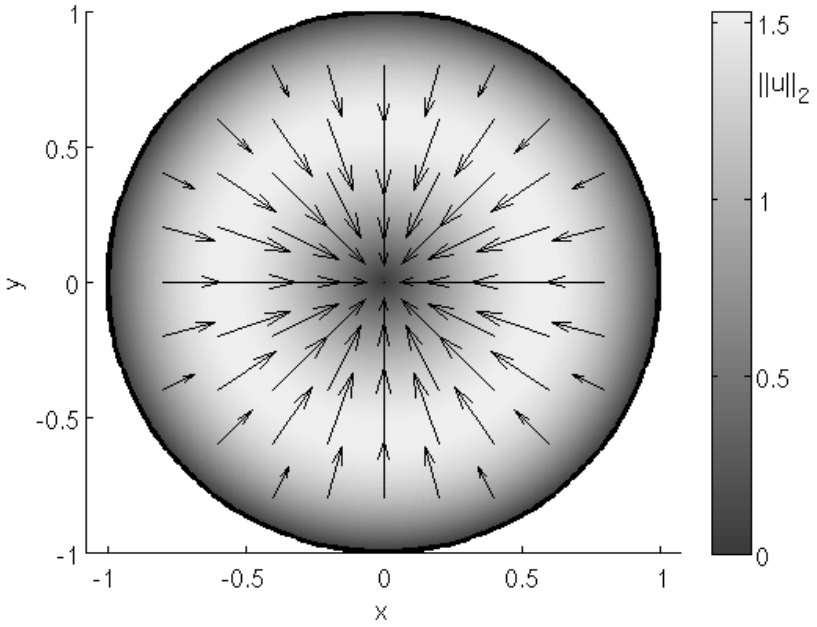
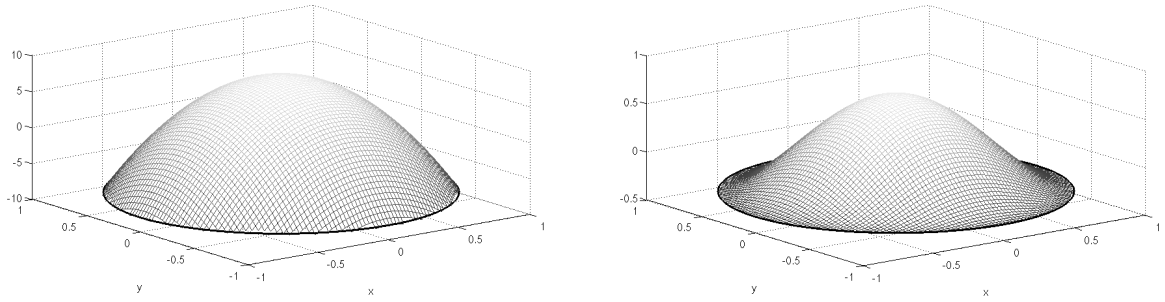


Figure 3.2: Exact solution (3.42): $\nabla p = -\mathbf{u}$



(a) div$\mathbf{u} = f$ (b) p

Figure 3.3: Exact solution (3.42)

3.3 Computational Results

In this section, the convergence result from Theorem 3.1 is illustrated on $\Omega = \{\mathbf{x} : \|x\|_2^2 \leq 1\}$. In order to find an exact solution, $\mathbf{u}$ is chosen such that $\mathbf{u} = 0$ on Γ, i.e.

$$\mathbf{u} = \bar{\mathbf{u}}(1 - \|\mathbf{x}\|_2^2) \ . \tag{3.41}$$

Note that then, the Neumann boundary conditions are trivially satified. In order to keep $\bar{\mathbf{u}}$ as simple as possible to get a gradient, let

$$\mathbf{u}(\mathbf{x}) = 4\mathbf{x}(1 - \|\mathbf{x}\|_2^2) \ , \ p(\mathbf{x}) = (1 - \|\mathbf{x}\|_2^2)^2 + \alpha \ . \tag{3.42}$$

As p has to solve the normalizing constraint, $\alpha = -\frac{1}{3}$ is setted. This leads to $f(\mathbf{x}) = 8 - 16\|\mathbf{x}\|_2^2$, as illustrated in figure 3.2 and 3.3. In figure 3.4, the results of the numerical computations are presented. On the top of the figure, the values of the least squares functional $\mathcal{F}_h(\mathbf{u}_h, p_h)$ are shown. Due to the fact that the exact solution is known, the values of the error norms $\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega_h}^2$ and $\|\nabla(p - p_h)\|_{0,\Omega_h}^2$ are presented as well in the bottom of the figure 3.4. The convergence rate that appears is in both norms inversely proportional to the number of unknowns N, that is proportional to h^{-2} in two dimensions. Due to the chosen finite-elements, this is the optimal order of convergence, and this confirms the predictions by Theorem 3.1.

In order to analyse if the optimal order of approximation is achievable with $(\mathbf{u}_h, p_h) \in RT_{1,\hat{\Gamma}}(\Omega) \times \dot{\mathcal{P}}_2(\Omega)$, the results for this combination of elements are plotted as well. The convergence rate is faster but the optimal convergence order 2 is not achieved. In order to get more indication, a second exemple is considered, where the right-hand side is given by $f(x_1, x_2) = \sin(x_1)$. The corresponding numerical results that are presented lemma in 3.5 confirm the order of convergence is no longer optimal for $(\mathbf{u}_h, p_h) \in RT_{1,\hat{\Gamma}}(\Omega) \times \dot{\mathcal{P}}_2(\Omega)$. This is due to the inaccurate representation of the boundary, and therefore a better approximation of the boundary and parametric elements on the domain with the interpolated boundary are needed to use higer-order elements. Their construction will be the topic of the next chapter.

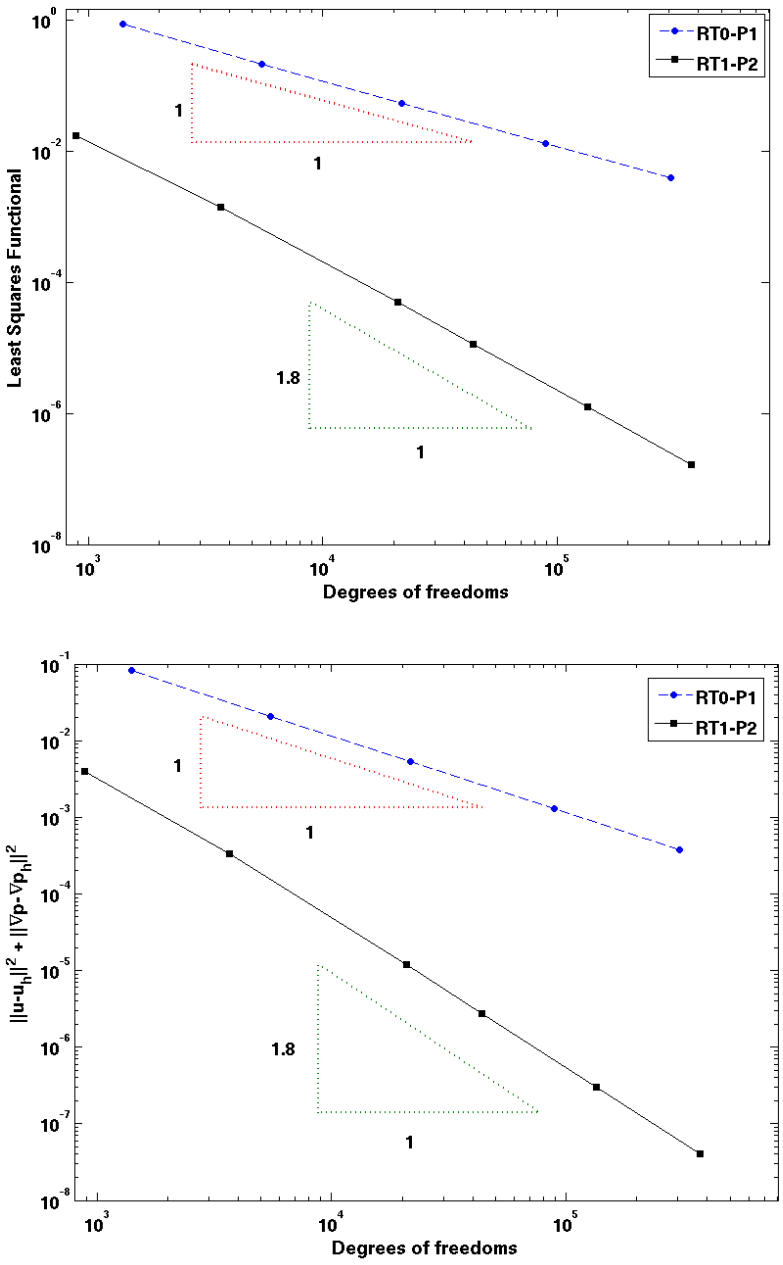


Figure 3.4: Convergence rates for minimizing $\mathcal{F}_h$ with $f(\mathbf{x}) = 8 - 16\|\mathbf{x}\|_2^2$

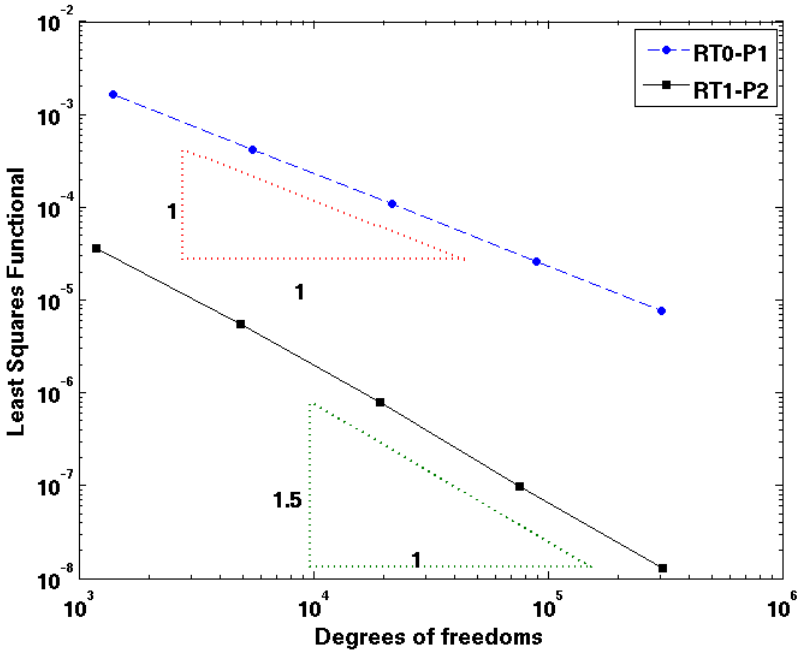


Figure 3.5: Convergence rates for minimizing $\mathcal{F}_h$ with $f(x_1, x_2) = \sin(x_1)$

Chapter 4

(Iso)parametric Finite Elements

The next of this work extends the study of the least squares method on domains with curved boundaries to the higher-order case. As shown at the end of the previous chapter, the polygonal approximation of the boundary does not lead to an optimal order of convergence in the higher-order case.

In the first section of this chapter, Ω is approximated by a domain Ω_h with piecewise polynomial boundary Γ_h. Note that for this reason, from now on $\Omega \subset \mathbb{R}^2$ is a bounded domain with a Lipschitz continuous and piecewise C^{k+2}-boundary $\Gamma = \partial\Omega$ for $k \geq 1$. Further, the mappings between Ω, Ω_h and $\hat{\Omega}$ are introduced and their properties are stated.

In the second section, finite-elements on the approximated domain Ω_h are introduced. These are isoparametric finite-elements of degree $k+1$ and parametric Raviart-Thomas elements of degree k. In particular, interpolation operators for these finite-element spaces are presented.

The use of the parametric Raviart-Thomas space permits the imposition of the Neumann boundary condition on Γ_h. Thus, the normal flux of the parametric Raviart-Thomas elements does not vanish on Γ and it has to be estimated. This is done in the third section.

4.1 Construction of the Approximated Domain Ω_h

The aim of this section is the construction of an approximated domain $\Omega_h = F_h(\hat{\Omega})$ with a piecewise polynomial mapping F_h such that the distance between $\Gamma_h = \partial\Omega_h$ and Γ is proportional to h^{k+2}. Therefore, the parametrisation $\boldsymbol{\gamma}$ (see Section 2.1) has to be interpolated by a polynomial parametrisation $\boldsymbol{\gamma}_h$ of degree $k+1$. This can be done using the polygonal parametrisation $\hat{\boldsymbol{\gamma}}$ and the one-dimensional Lagrange interpolation. On each boundary segment $\hat{\Gamma}_i$ let $\{\hat{\boldsymbol{\gamma}}_{i,j}\}_{j=0}^{k+1}$ denote the $k+2$ equidistant points:

$$\hat{\boldsymbol{\gamma}}_{i,j} = \left(1 - \frac{j}{k+1}\right)\boldsymbol{\gamma}(t_{i-1}) + \frac{j}{k+1}\boldsymbol{\gamma}(t_i) \ . \tag{4.1}$$

Using the mapping $\hat{\boldsymbol{\zeta}}$, one obtains the set of interpolation points $\{\boldsymbol{\gamma}_{i,j}\}_{j=0}^{k+1}$ with

$$\boldsymbol{\gamma}_{i,j} = \hat{\boldsymbol{\zeta}}(\hat{\boldsymbol{\gamma}}_{i,j}) \ . \tag{4.2}$$

The approximation Γ_h is then parametrized by

$$\Gamma_h = \bigcup_{i=1}^{N} \left\{ \boldsymbol{\gamma}_{h,i}(t) = \sum_{j=0}^{k+1} \boldsymbol{\gamma}_{i,j}\mathscr{L}_{i,j}(t) \ : \ 1 \leq i \leq N \text{ and } t_{i-1} \leq t \leq t_i \right\} \tag{4.3}$$

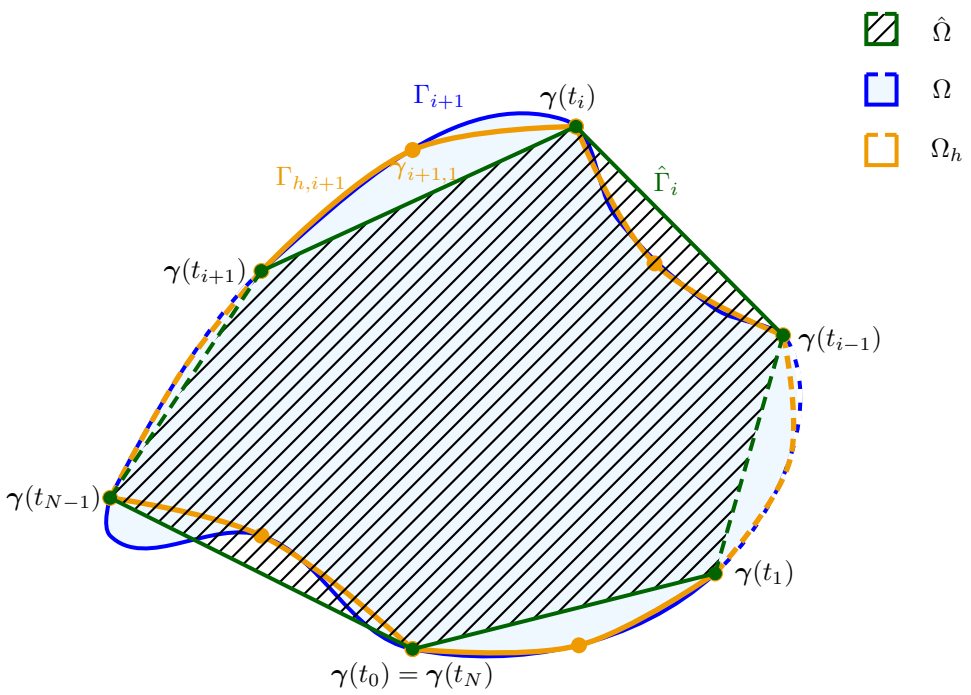


Figure 4.1: Quadratic interpolation of the boundary

of Lagrange basis polynomials with the Lagrange basis polynomial

$$\mathscr{L}_{i,j}(t) := \prod_{\substack{0 \le m \le k+1 \\ m \neq j}} \frac{t - t_{i,m}}{t_{i,j} - t_{i,m}} \text{ where } t_{i,j} = t_{i-1} + \frac{j}{k+1} t_i \ . \tag{4.4}$$

Such an approximation Γ_h is depicted in Figure 4.1. Note that the length of each boundary segement $\Gamma_{h,i} = \{\boldsymbol{\gamma}_h(t) \ : \ t_{i-1} \le t \le t_i\}$ is smaller than h due to the property of $\hat{\Gamma}$ in Section 2.1. Further, assume that all points where Γ is not C^{k+2} are also vertices of the triangulation such that Γ_i is a C^{k+2} curve and thus from the classical interpolation theory it holds

$$\|\boldsymbol{\gamma} - \boldsymbol{\gamma}_h\|_{l,\infty,[t_0,t_N]} \lesssim h^{k+2} |D^2 \boldsymbol{\gamma}|_{\infty,[t_0,t_N]} \qquad \forall l \le k+2 \tag{4.5}$$

and if Ω_h denote the domain bounded by Γ_h,

$$|\Omega \backslash \Omega_h| \eqsim |\Omega_h \backslash \Omega| \eqsim h^{k+2}. \tag{4.6}$$

As in Section 2.1, a triangulation $\mathcal{T}_h$ with curved triangles can be defined from $\hat{\mathcal{T}}_h$ by replacing each straight boundary segment by the approximated one: $\mathcal{T}_h$ consists then on $\bar{N}$ triangles and for $i \le N$, T_i denotes the approximated triangle corresponding to $\hat{T}_i$ (and to $\widetilde{T}_i$), i.e. T_i is the shape with the same vertices as $\hat{T}_i$ and whose boundary consits of the segments $\partial \hat{T}_i \backslash \hat{\Gamma}_i$ and the arc Γ_i. For $i > N$, T_i is identical to $\hat{T}_i$. Thus, Ω_h is completely covered by $\mathcal{T}_h$, as illustrated in Figure 4.2.

For the construction of the mapping $F_h : \hat{\Omega} \to \Omega_h$ the same way as in Section 2.3 is used, replacing Ω by Ω_h. To this end, first define $\boldsymbol{\zeta} = \boldsymbol{\gamma}_h \circ \hat{\boldsymbol{\gamma}}^{-1}$, such that $\boldsymbol{\zeta}$ maps $\hat{\Gamma}$ on Γ_h. Then, the

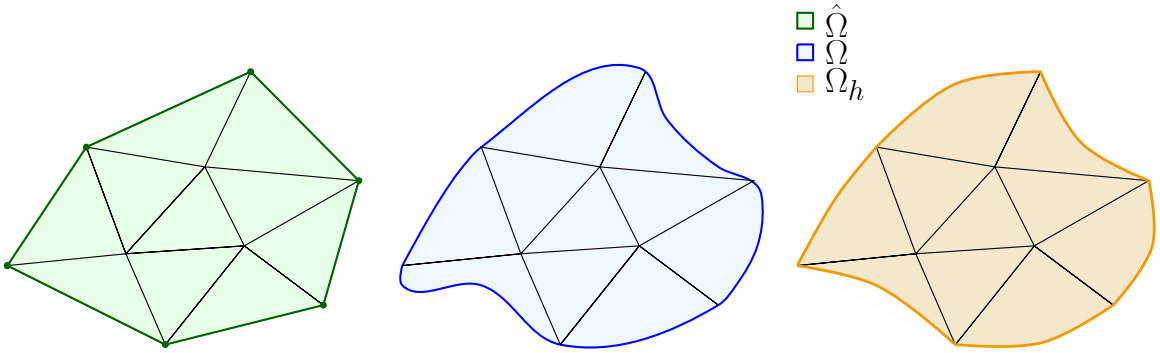


Figure 4.2: Triangulations of Ω, $\hat{\Omega}$ and Ω_h

mapping $F_{h,i} : \hat{T}_i \to T_i$ can be defined as

$$F_{h,i}(\mathbf{x}) = \mathbf{Y}_{\delta(\mathbf{x})}(\boldsymbol{\zeta}(\mathbf{Z}(\mathbf{x}))). \tag{4.7}$$

with the functions $\mathbf{Y}$ and $\mathbf{Z}$ defined in Section 2.3. Setting $F_{h,i} = \mathrm{id}$ for the interior elements leads to the global definition of F_h. Similarly to the derivation of the properites of $\hat{\boldsymbol{\Phi}}_h$ in Section 2.3, note that due to its construction, $F_{h,i}$ is the identity mapping on the two edges of $\hat{T}_i$ which have one single common point with $\hat{\Gamma}$ and it holds

$$F_h \circ \hat{\boldsymbol{\gamma}} = \boldsymbol{\gamma}_h \tag{4.8}$$

and

$$\begin{aligned} F_{h,i}(\mathbf{x}) - \mathbf{x} &= \mathbf{Y}_{\delta(\mathbf{x})}(\boldsymbol{\zeta}(\mathbf{Z}(\mathbf{x}))) - \mathbf{Y}_{\delta(\mathbf{x})}(\mathbf{Z}(\mathbf{x})) \\ &= \mathbf{Y}_{\delta(\mathbf{x})}((\boldsymbol{\zeta} - \mathrm{id})(\mathbf{Z}(\mathbf{x}))) \\ &= \mathbf{Y}_{\delta(\mathbf{x})}((\boldsymbol{\gamma}_h \circ \hat{\boldsymbol{\gamma}}^{-1} - \hat{\boldsymbol{\gamma}} \circ \hat{\boldsymbol{\gamma}}^{-1})(\mathbf{Z}(\mathbf{x}))) \\ &= \mathbf{Y}_{\delta(\mathbf{x})}(((\boldsymbol{\gamma}_h - \hat{\boldsymbol{\gamma}}) \circ \hat{\boldsymbol{\gamma}}^{-1})(\mathbf{Z}(\mathbf{x}))) \ . \end{aligned} \tag{4.9}$$

Combining this with (4.5) leads to the following theorem (see [27] Theorem 1)

Theorem 4.1. *With the previous assumptions and $\hat{T}_i \in \hat{\mathcal{T}}_h$, $F_{h,i}$ is a $\mathcal{C}^{k+2}$ diffeomorphism, polynomial of degree $k+1$ and invertible in a neighborhood of $\hat{\Omega}$. Moreover, for a positive integer s with $s \leq k+2$, it holds*

$$\|F_h\|_{W^s_\infty(\hat{T}_i)} \lesssim h^s \tag{4.10a}$$

$$\|F_h^{-1}\|_{W^s_\infty(\hat{T}_i)} \lesssim h^{-s} \ . \tag{4.10b}$$

Note that the invertibility of $F_{h,i}$ in a neighborhood of $\hat{\Omega}$ can be used by the definition of the finite-elements to extend it from Ω_h to Ω for small enough h. Further, an auxiliary mapping $\boldsymbol{\Phi}_h : \Omega_h \to \Omega$ can be given by

$$\boldsymbol{\Phi}_h = \hat{\boldsymbol{\Phi}}_h \circ F_h^{-1} \ . \tag{4.11}$$

Combining the properties from $\hat{\boldsymbol{\Phi}}_h$ and F_h leads to the following theorem (see [27] Paragraph 5):

Theorem 4.2. *With the previous assumptions and $\hat{T}_i \in \hat{\mathcal{T}}_h$, $\mathbf{\Phi}_{h,i}$ is a $\mathcal{C}^{k+2}$ diffeomorphism and for a positive integer with $s \leq k+2$, it holds*

$$\|\mathbf{\Phi} - id\|_{W^s_\infty(T_i)} \lesssim h^{k+2-s} \tag{4.12a}$$

$$\text{and } \|\det(J_{\mathbf{\Phi}_{i,h}}) - 1\|_{L^\infty(T_i)} \lesssim h^{k+1} \tag{4.12b}$$

Further, the mapping $\mathbf{\Phi}_{h,i} : T_i \to \widetilde{T}_i$ is injective and $\mathbf{\Psi}_{h,i} = \mathbf{\Phi}_{h,i}^{-1}$ satisfies

$$\|\mathbf{\Psi} - id\|_{W^s_\infty(\widetilde{T}_i)} \lesssim h^{k+2-s} \quad \forall s \leq k+2 \tag{4.13a}$$

$$\text{and } \|\det(J_{\mathbf{\Psi}_{i,h}}) - 1\|_{L^\infty(\widetilde{T}_i)} \lesssim h^{k+1} \tag{4.13b}$$

Proof: Only the first step is given here, for more details, see [27]. Therefore, note that $\mathbf{\Phi}_{h,i}$ is the identity mapping on the two edges of T_i which have one single common point with Γ_h. Further, due to its construction it holds

$$\mathbf{\Phi}_h \circ \boldsymbol{\gamma}_h = \boldsymbol{\gamma} \tag{4.14}$$

and

$$\begin{aligned}
\mathbf{\Phi}_{h,i}(\mathbf{x}) - \mathbf{x} &= \mathbf{Y}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}(\hat{\boldsymbol{\zeta}}(\boldsymbol{\zeta}^{-1}(\mathbf{Y}^{-1}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}))) - \mathbf{Y}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}(\mathbf{Y}^{-1}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}) \\
&= \mathbf{Y}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}((\boldsymbol{\gamma} \circ \boldsymbol{\gamma}_h^{-1} - \boldsymbol{\gamma}_h \circ \boldsymbol{\gamma}_h^{-1})(\mathbf{Y}^{-1}_{\delta(F_{h,i}^{-1}(\mathbf{x}))})) \\
&= \mathbf{Y}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}(((\boldsymbol{\gamma} - \boldsymbol{\gamma}_h) \circ \boldsymbol{\gamma}_h^{-1})(\mathbf{Y}^{-1}_{\delta(F_{h,i}^{-1}(\mathbf{x}))}))
\end{aligned}$$

such that the further steps consist in using (4.5) and the smoothness properties of $\mathbf{Y}$ and $\boldsymbol{\gamma}_h^{-1}$. □

4.2 (Iso)parametric Spaces

In order to define finite-elements on Ω_h, the parametric Raviart-Thomas elements of degree k on those triangles containing a curved boundary edge are introduced. It is defined as

$$\mathbf{V}_h^k|_{\Omega_h} = \{\mathbf{v}_h : \Omega_h \to \mathbb{R}^2 : \mathbf{v}_h = \left(\frac{1}{\det J_{F_h}} J_{F_h} \hat{\mathbf{v}}_h\right) \circ F_h^{-1} \text{ with } \hat{\mathbf{v}}_h \in RT_{k,\hat{\Gamma}}(\hat{\Omega}, \mathcal{T}_h)\}\,. \tag{4.15}$$

Note that as in Section 2.2, these elements can be extended to Ω using the polynomial representation and the fact that $F_h : \hat{\Omega} \to \Omega_h$ is invertible in a neighborhood of $\hat{\Omega}$ for a small enough h. This leads to

$$\mathbf{V}_h^k = \{\mathbf{v}_h : \Omega \to \mathbb{R}^2 : \mathbf{v}_h = \left(\frac{1}{\det J_{F_h}} J_{F_h} \hat{\mathbf{v}}_h\right) \circ F_h^{-1} \text{ with } \hat{\mathbf{v}}_h \in RT_{k,\hat{\Gamma}}(\hat{\Omega}, \mathcal{T}_h)\}\,. \tag{4.16}$$

Due to Theorem 1.8 and as the normal flux of $\hat{\mathbf{v}}_h \in RT_{k,\hat{\Gamma}}(\hat{\Omega}, \mathcal{T}_h)$ vanishes on $\hat{\Gamma}$, the normal flux of $\mathbf{v}_h \in \mathbf{V}_h^k$ vanishes on Γ_h and it holds

$$\mathbf{V}_h^k \subset H_{\Gamma_h}(\text{div}, \Omega_h). \tag{4.17}$$

In order to approximate the pressure, the standard scalar isoparametric finite-elements is used:

$$Q_h^{k+1}|_{\Omega_h} = \{q_h : \Omega_h \to \mathbb{R}^2 : q_h = \hat{q}_h \circ F_h^{-1} \text{ with } \hat{q}_h \in \mathcal{P}_{k+1}(\hat{\Omega})\}\,. \tag{4.18}$$

As for the Raviart-Thomas elements, the polynomial representation are used to extend the elements on Ω. Further, an implementable normalizing constraint is added into the space to get the finite-element space used to approximate the pressure:

$$\acute{Q}_h^{k+1} = \{q_h : \Omega \to \mathbb{R}^2 : q_h = \hat{q}_h \circ F_h^{-1} \text{ with } \hat{q}_h \in \mathcal{P}_{k+1}(\hat{\Omega}) \text{ and } (q_h, 1)_{0,\Omega_h} = 0\} \,. \tag{4.19}$$

However, similar to Section 2.2 and in order to simplify the analysis in the next part of this work, the following space is used for the analysis

$$\dot{Q}_h^{k+1} = \{q_h : \Omega \to \mathbb{R}^2 : q_h = \hat{q}_h \circ F_h^{-1} \text{ with } \hat{q}_h \in \mathcal{P}_{k+1}(\hat{\Omega}) \text{ and } (q_h, 1)_{0,\Omega} = 0\} \,, \tag{4.20}$$

since two solutions $\acute{p_h} \in \acute{Q}_h^{k+1}$ and $p_h \in \dot{Q}_h^{k+1}$ of the problem (1.36) where the space for p_h is replaced with the corresponding one, only differ by a constant on the order of h^{k+1}: it holds $p_h = \acute{p_h} - C$ with

$$C = \frac{(\acute{p_h}, 1)_{0,\Omega}}{(1,1)_{0,\Omega}} \tag{4.21}$$

since $p_h - \acute{p_h}$ and thus

$$p_h - \acute{p_h} = \frac{(p_h - \acute{p_h}, 1)_{0,\Omega}}{(1,1)_{0,\Omega}} = -\frac{(\acute{p_h}, 1)_{0,\Omega}}{(1,1)_{0,\Omega}} \,. \tag{4.22}$$

Further, it holds (see [8])

$$\begin{aligned}(\acute{p_h}, 1)_{0,\Omega} &= (\acute{p_h}, 1)_{0,\Omega} - (\acute{p_h}, 1)_{0,\Omega_h} = \int_\Omega \left(\acute{p_h} - (\acute{p_h} \circ \mathbf{\Psi}_h)(\det J_{\mathbf{\Psi}_h})\right) \, d\mathbf{x} \\ &= \int_\Omega \acute{p_h}(1 - \det J_{\mathbf{\Psi}_h}) \, dx + \int_\Omega (\acute{p_h} - (\acute{p_h} \circ \mathbf{\Psi}_h)) \, (\det J_{\mathbf{\Psi}_h}) \, d\mathbf{x} \,.\end{aligned}$$

Using (4.13b) leads to

$$\begin{aligned}(\acute{p_h}, 1)_{0,\Omega} &\lesssim h^{k+1} \|\acute{p_h}\|_{0,\Omega} + \int_\Omega \int_0^1 \left| \frac{d}{ds} \acute{p_h}(\mathbf{x} + s(\mathbf{\Psi}_h(\mathbf{x}) - \mathbf{x})) \right| \, ds \, d\mathbf{x} \\ &\leq h^{k+1} \|\acute{p_h}\|_{0,\Omega} + \int_\Omega \int_0^1 \left| \nabla \acute{p_h}(\mathbf{x} + s(\mathbf{\Psi}_h(\mathbf{x}) - \mathbf{x})) \right| \, ds \, \left| \mathbf{\Psi}_h(\mathbf{x}) - \mathbf{x} \right| \, d\mathbf{x} \\ &\leq h^{k+1} \|\acute{p_h}\|_{0,\Omega} + \|\nabla \acute{p_h}\|_{L^\infty(\Omega)} \|\mathbf{\Psi}_h - \mathrm{id}\|_{0,\Omega} \,.\end{aligned}$$

Using (4.13a) implies

$$(\acute{p_h}, 1)_{0,\Omega} \lesssim h^{k+1} \left(\|\acute{p_h}\|_{0,\Omega} + h \|\nabla \acute{p_h}\|_{L^\infty(\Omega)} \right) .$$

At this point, a scaling argument is needed:

$$\begin{aligned}\|\nabla \acute{p_h}\|_{L^\infty(\Omega)} &= \sum_{i=1}^{\bar{N}} \|\nabla \acute{p_h}\|_{L^\infty(\widetilde{T}_i)} = \sum_{i=1}^{\bar{N}} \|(\nabla \acute{p_h}) \circ \mathbf{F}_{\hat{T}_i, \mathrm{ref}}\|_{L^\infty(T_{ref})} \\ &\lesssim \sum_{i=1}^{\bar{N}} \|(\nabla \acute{p_h}) \circ \mathbf{F}_{\hat{T}_i, \mathrm{ref}}\|_{0,T_{ref}} \lesssim \sum_{i=1}^{\bar{N}} h_i^{-1} \|\nabla \acute{p_h}\|_{0,\widetilde{T}_i} \\ &\lesssim h^{-1} \|\nabla \acute{p_h}\|_{0,\Omega} \,.\end{aligned}$$

This leads to

$$(\acute{p_h}, 1)_{0,\Omega} \lesssim h^{k+1} \left(\|\acute{p_h}\|_{0,\Omega} + \|\nabla \acute{p_h}\|_{0,\Omega} \right) \lesssim h^{k+1} \|\acute{p_h}\|_{1,\Omega} \,.$$

such that

$$C \lesssim h^{k+1} \|\hat{p}_h\|_{1,\Omega} \ . \tag{4.23}$$

In order to apply the finite-element theory developed in Chapter 1 to these (iso)parametric spaces, interpolation operators are needed. On one hand for the isoparametric approximation of $H^1(\Omega)$ define

$$\mathcal{Q}_h q = (\mathcal{I}_h(q \circ F_h)) \circ F_h^{-1} \ . \tag{4.24}$$

Then, for $s = 0$ or $s = 1$ the classical isoparametric theory (see [14], Theorem 4.7.3.) states

$$\|q - \mathcal{Q}_h q\|_{W_p^s}(\Omega_h) \lesssim h^{k+2-s} \|q\|_{W_p^{k+2}} \tag{4.25a}$$

$$\|q - \mathcal{Q}_h q\|_{W_\infty^s}(\Omega_h) \lesssim h^{k+2-s-2/p} \|q\|_{W_p^{k+2}} \ . \tag{4.25b}$$

On the other hand, for the parametric Raviart-Thomas finite-element space define

$$(\mathcal{R}_h \mathbf{v}) = \left(\frac{1}{\det J_{F_h}} J_{F_h} (\hat{\mathcal{R}}_h \hat{\mathbf{v}}) \right) \circ F_h^{-1} \ , \tag{4.26}$$

where $\hat{v} = (\det J_{F_h}) J_{F_h}^{-1} (\mathbf{v} \circ F_h)$. Note that due to the definition of $\hat{\mathcal{R}}_h$ and (1.8), $\mathcal{R}_h$ maps $\mathbf{v} \in H_{\Gamma_h}(\mathrm{div}, \Omega_h)$ on $\mathbf{V}_h^k$. Further, it holds

$$\mathbf{v} - (\mathcal{R}_h \mathbf{v}) = \left(\frac{1}{\det J_{F_h}} J_{F_h} (\hat{\mathbf{v}} - \hat{\mathcal{R}}_h \hat{\mathbf{v}}) \right) \circ F_h^{-1} \tag{4.27a}$$

$$\text{and} \quad \mathrm{div}\ (\mathbf{v} - (\mathcal{R}_h \mathbf{v})) = \frac{1}{\det J_{F_h}} \mathrm{div} \left(\left(\frac{1}{\det J_{F_h}} J_{F_h} (\hat{\mathbf{v}} - \hat{\mathcal{R}}_h \hat{\mathbf{v}}) \right) \circ F_h^{-1} \right) \tag{4.27b}$$

which leads to

$$\|\mathbf{v} - \mathcal{R}_h \mathbf{v}\|_{0,\Omega_h} \leq \|(\det J_{F_h})^{-1}\|_{L^\infty(\hat{\Omega})} \|J_{F_h}\|_{L^\infty(\hat{\Omega})} \|\hat{\mathbf{v}} - \hat{\mathcal{R}}_h \hat{\mathbf{v}}\|_{0,\hat{\Omega}} \tag{4.28a}$$

$$\text{and} \quad \|\mathrm{div}\ (\mathbf{v} - \mathcal{R}_h \mathbf{v})\|_{0,\Omega_h} \leq \|(\det J_{F_h})^{-1}\|_{L^\infty(\hat{\Omega})} \|J_{F_h}\|_{L^\infty(\hat{\Omega})} \|\mathrm{div}\ (\hat{\mathbf{v}} - \hat{\mathcal{R}}_h \hat{\mathbf{v}})\|_{0,\hat{\Omega}} \ . \tag{4.28b}$$

Using the standard interpolation estimates (1.59) and the properties of F_h stated in (4.10), this turns into

$$\|\mathbf{v} - \mathcal{R}_h \mathbf{v}\|_{0,\Omega_h} \lesssim \|\hat{\mathbf{v}} - \hat{\mathcal{R}}_h \hat{\mathbf{v}}\|_{0,\hat{\Omega}} \lesssim h^{k+1} \|\hat{\mathbf{v}}\|_{k+1,\hat{\Omega}} \tag{4.29a}$$

$$\text{and} \quad \|\mathrm{div}\ (\mathbf{v} - \mathcal{R}_h \mathbf{v})\|_{0,\Omega_h} \lesssim \|\mathrm{div}\ (\hat{\mathbf{v}} - \hat{\mathcal{R}}_h \hat{\mathbf{v}})\|_{0,\hat{\Omega}} \lesssim h^{k+1} \|\mathrm{div}\ \hat{\mathbf{v}}\|_{k+1,\hat{\Omega}}, \tag{4.29b}$$

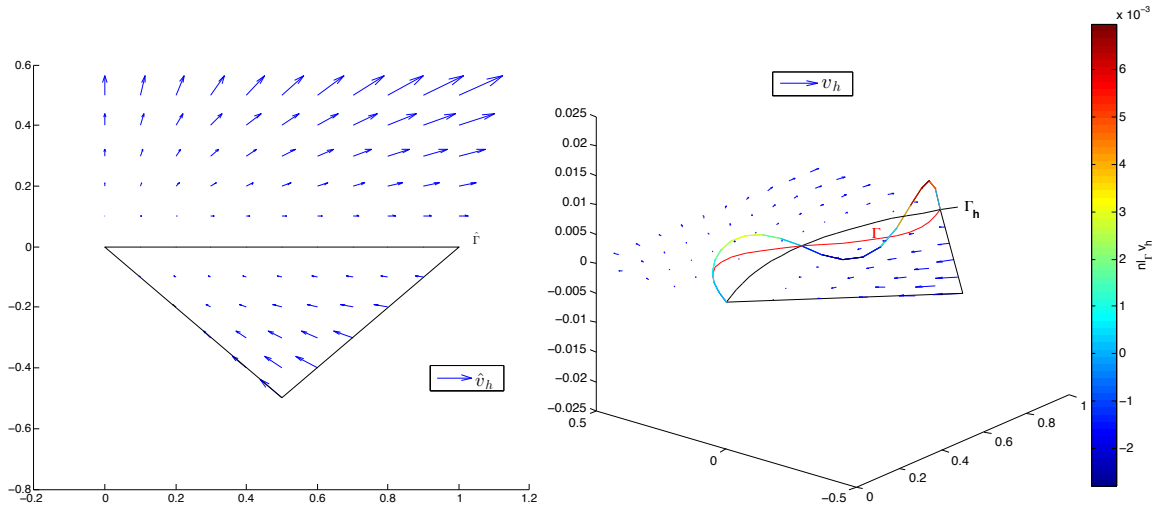
for all $\mathbf{v} \in H^{k+1}(\Omega_h)^2$ with $\mathrm{div}\, \mathbf{v} \in H^{k+1}(\Omega_h)$. It remains to transform the right terms of these inequalities on Ω_h. This can be done using repeatedly the chain rule and (4.10) to obtain

$$\|\mathbf{v} - \mathcal{R}_h \mathbf{v}\|_{0,\Omega_h} \lesssim h^{k+1} \|\mathbf{v}\|_{k+1,\Omega} \tag{4.30a}$$

$$\text{and} \quad \|\mathrm{div}\ (\mathbf{v} - \mathcal{R}_h \mathbf{v})\|_{0,\Omega_h} \lesssim h^{k+1} \|\mathrm{div}\ \mathbf{v}\|_{k+1,\Omega} \ . \tag{4.30b}$$

4.3 An Estimate for the Normal Flux on Interpolated Boundaries

The normal flux of functions belonging to $\mathbf{V}_h^k$ vanishes only on Γ_h and not on Γ (see Figure 4.3), such that $\mathbf{V}_h^k$ is not a subspace of $H_\Gamma(\mathrm{div}, \Omega)$. Therefore, an estimate for the approximated normal flux $\mathbf{n} \cdot \mathbf{v}_h$ on Γ for $\mathbf{v}_h \in \mathbf{V}_h^k$ is needed. This is the statement of the following theorem (see [8]).



(a) $\hat{\mathbf{v}}_h \cdot \mathbf{n} = 0$ on $\hat{\Gamma}$ (b) $\mathbf{v}_h \cdot \mathbf{n} = 0$ on Γ_h and $\mathbf{v}_h \cdot \mathbf{n} \neq 0$ on Γ

Figure 4.3: Normal flux on interpolated boundaries

Theorem 4.3. *Let Ω having the properties described in Section 4.1. In particular, assume that Γ is a piecewise C^{k+2} curve, $k \geq 0$. Then,*

$$|\langle \mathbf{n} \cdot \mathbf{v}_h, q \rangle_{0,\Gamma}| \lesssim h^{k+1} \left(\|\mathbf{v}_h\|_{0,\Omega}^2 + \|\mathrm{div}\, \mathbf{v}_h\|_{0,\Omega}^2 \right)^{\frac{1}{2}} \|q\|_{\frac{1}{2},\Gamma} \tag{4.31}$$

holds for all $\mathbf{v}_h \in \mathbf{V}_h^k$ and $q \in H^{\frac{1}{2}}(\Gamma)$.

Proof: Consider the curved boundary edge Γ_i. In order to simplify the notation, the isometry $\mathbf{D}_i$ from Section 2.4 is used in combination with $F_{h,i}^{-1}$ to map Γ_h on $[0, h_i]$, as represented in Figure 4.4. The image of the curved edge Γ_i under $\mathbf{D}_i \circ F_{h,i}^{-1}$ can then be given as the graph of a function η:

$$\left(\mathbf{D}_i \circ F_{h,i}^{-1} \right) (\Gamma_i) = \Gamma_{\mathbf{D},i} = \left\{ \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix} : 0 \leq \xi \leq h_i \right\} . \tag{4.32}$$

where

$$\begin{aligned} \eta : [0, h_i] &\to \mathbb{R}^2 \\ \xi &\mapsto \left(\mathbf{D}_i \circ F_{h,i}^{-1} \circ \hat{\boldsymbol{\zeta}} \circ \mathbf{D}_i^{-1} \right) ((0, \xi)) . \end{aligned} \tag{4.33}$$

As $\hat{\boldsymbol{\zeta}}$ is a C^{k+2}-diffeomorphismus, η is a C^{k+2}-function. Thus, due to

$$\eta \left(j \frac{h_i}{k+1} \right) = 0 \tag{4.34}$$

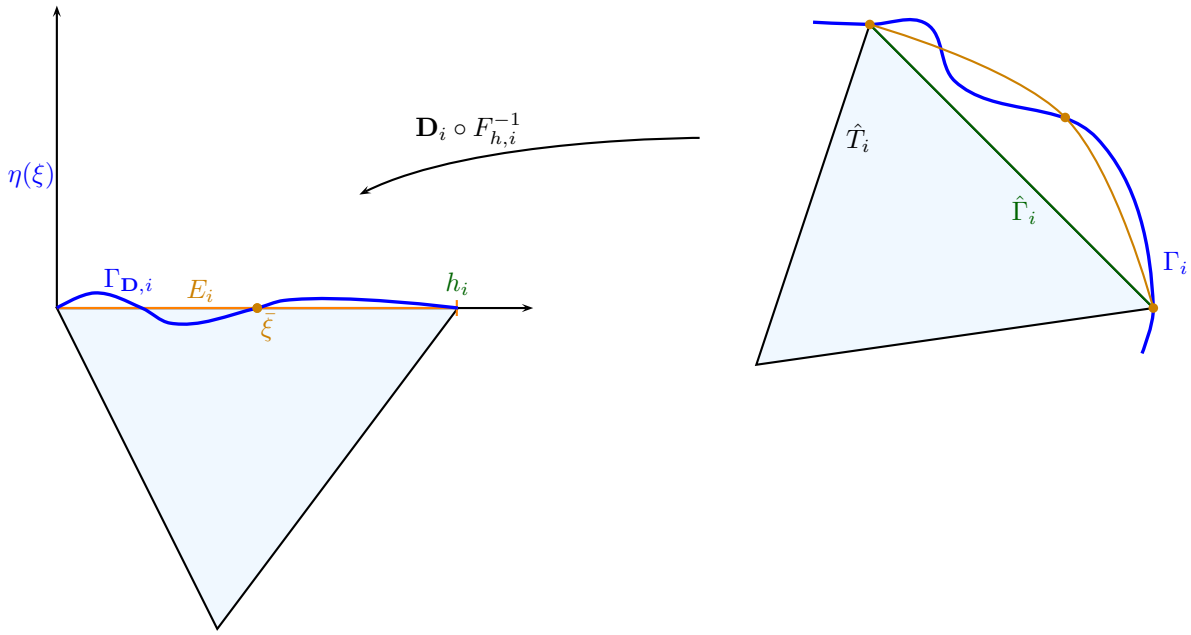


Figure 4.4: Mapping $\mathbf{D}_i \circ F_{h,i}^{-1}$

for $0 \leq j \leq k+1$, it holds

$$|\eta(\xi)| \lesssim h_i^{k+2} \tag{4.35a}$$

$$\text{and } |\eta'(\xi)| \lesssim h_i^{k+1} \tag{4.35b}$$

for all $\xi \in [0, h_i]$. Now, let be $\mathbf{v}_h \in \mathbf{V}_h^k$. Racall that then

$$\mathbf{v}_h = \left(\frac{1}{\det J_{F_h}} J_{F_h} \hat{\mathbf{v}}_h \right) \circ F_h^{-1} \tag{4.36}$$

with $\hat{\mathbf{v}}_h \in RT_{k,\hat{\Gamma}}(\Omega)$. Similarly to the proof of Theorem 2.1, define $\hat{\mathbf{w}}_{h,i} = \hat{\mathbf{v}}_{h,i} \circ \mathbf{D}_i^{-1}$. Due to $\det(J_{\mathbf{D}_i}) = 1$ it holds

$$\|\hat{\mathbf{w}}_{h,i}\|_{0,\mathbf{D}_i(\hat{T}_i)} = \|\hat{\mathbf{v}}_{h,i}\|_{0,\hat{T}_i} \tag{4.37a}$$

$$\hat{\mathbf{w}}_{h,i} \cdot \mathbf{n} = 0 \text{ on } E_i = \{(\xi, 0) \ : \ 0 \leq \xi \leq h_i\} \tag{4.37b}$$

$$\text{and } \int_{\Gamma_i} (\hat{\mathbf{v}}_{h,i} \cdot \mathbf{n}) q \, \mathrm{d}s = \int_0^{h_i} (\hat{\mathbf{w}}_{h,i} \cdot \mathbf{n})(q \circ \mathbf{D}_i^{-1}) \, \mathrm{d}\xi \tag{4.37c}$$

for $q \in L^2(\hat{\Gamma}_i)$. Then, the special form of $\mathbf{w}_{h,i} \in RT_{k,E_i}(\mathbf{D}_i(\widetilde{T}_i))$ can be used as $\mathbf{w}_{h,i}(\mathbf{x}) \cdot \mathbf{n} = 0$ for all $\mathbf{x} \in E_i$ implies

$$\mathbf{w}_{h,i}(\xi, 0) \cdot \mathbf{n} = \left(\boldsymbol{\alpha}_k(\xi, 0) + \beta_k \begin{pmatrix} \xi \\ 0 \end{pmatrix} \right) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \quad \forall \xi \in [0, h_i] \tag{4.38}$$

and thus

$$[\boldsymbol{\alpha}(\xi,0)]_2 = \sum_{j_1=0}^{k}\sum_{j_2=0}^{j_1} c_{j_1,j_2}\xi^{j_1-j_2}0^{j_2} = 0 \quad \forall \xi \in [0,h_i]. \tag{4.39}$$

This means that c_{j_1,j_2} has to vanish for $j_2 = 0$ such that

$$[\boldsymbol{\alpha}(x,y)]_2 = \sum_{j_1=0}^{k}\sum_{j_2=1}^{j_1} c_{j_1,j_2}x^{j_1-j_2}y^{j_2} = y\sum_{j_1=0}^{k}\sum_{j_2=1}^{j_1} c_{j_1,j_2}x^{j_1-j_2}y^{j_2-1} = y\bar{\alpha}_k(x,y)\ . \tag{4.40}$$

Denoting $\alpha_k(x,y) = [\boldsymbol{\alpha}(x,y)]_1$ leads to

$$\hat{\mathbf{w}}_{h,i}(\mathbf{x}) = \begin{pmatrix}\alpha_k(\mathbf{x})\\ y\bar{\alpha}_k(\mathbf{x})\end{pmatrix} + \mathbf{x}\beta_k(\mathbf{x})\ . \tag{4.41}$$

This implies that

$$\hat{\mathbf{w}}_{h,i}(\xi,0) = \begin{pmatrix}\alpha_k(\xi,0)+\xi\beta_k(\xi,0)\\ 0\end{pmatrix} \text{ on } E_i \tag{4.42a}$$

$$\text{and } \mathbf{n}\cdot\hat{\mathbf{w}}_{h,i} = \frac{1}{(1+\eta'(\xi)^2)^{\frac{1}{2}}}\begin{pmatrix}-\eta'(\xi)\\ 1\end{pmatrix}\cdot\begin{pmatrix}\alpha_k(\xi,\eta(\xi))+\xi\ \beta_k(\xi,\eta(\xi))\\ \eta(\xi)\ (\bar{\alpha}_k(\xi,\eta(\xi)+\beta_k(\xi,\eta(\xi)))\end{pmatrix} \text{ on } \Gamma_{\mathbf{D},i}\ . \tag{4.42b}$$

Then, first note that using (4.35b) leads to

$$\begin{aligned}\left(\int_0^{h_i}(\eta'(\xi))^2(\alpha_k(\xi,\eta(\xi))+\xi\ \beta_k(\xi,\eta(\xi)))^2 d\xi\right)^{\frac{1}{2}} &\lesssim h_i^{k+1}\left(\int_0^{h_i}([\hat{\mathbf{w}}_{h,i}]_1(\xi,\eta(\xi)))^2 d\xi\right)^{\frac{1}{2}}\\ &\lesssim h_i^{k+1}\|[\hat{\mathbf{w}}_{h,i}]_1\|_{0,\Gamma_{\mathbf{D},i}}\end{aligned}$$

and thus to

$$\left(\int_0^{h_i}(\eta'(\xi))^2(\alpha_k(\xi,\eta(\xi))+\xi\ \beta_k(\xi,\eta(\xi)))^2 d\xi\right)^{\frac{1}{2}} \lesssim h_i^{k+\frac{1}{2}}\|[\hat{\mathbf{w}}_{h,i}]_1\|_{0,\mathbf{D}_i(\hat{T}_i)} \tag{4.43}$$

using the scaling argument (2.41). Further, due to the Leibniz rule for the derivatives of products it holds

$$\begin{aligned}\frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}(x,y)]_2 &= \frac{d^{l+1}}{dy^{l+1}}\left(y\left(\bar{\alpha}_k(x,y)+\beta_k(x,y)\right)\right)\\ &= \sum_{j=0}^{l+1}\binom{l+1}{j}\frac{d^j}{dy^j}y\frac{d^{l+1-j}}{dy^{l+1-j}}\left(\bar{\alpha}_k(x,y)+\beta_k(x,y)\right)\\ &= \sum_{j=0}^{1}\binom{l+1}{j}\frac{d^j}{dy^j}y\frac{d^{l+1-j}}{dy^{l+1-j}}\left(\bar{\alpha}_k(x,y)+\beta_k(x,y)\right)\\ &= y\frac{d^{l+1}}{dy^{l+1}}\left(\bar{\alpha}_k(x,y)+\beta_k(x,y)\right)+(l+1)\frac{d^l}{dy^l}\left(\bar{\alpha}_k(x,y)+\beta_k(x,y)\right)\end{aligned} \tag{4.44}$$

for any non-negative integer $l \le k$, and thus

$$\frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)} = (l+1)\frac{d^l}{dy^l}\left(\bar{\alpha}_k(x,y)+\beta_k(x,y)\right)\Big|_{(\xi,0)}\ .$$

Combining this with the fact that $y \mapsto (\bar{\alpha}_k(x,y) + \beta_k(x,y))$ is a polynom from grad k and thus that it can be written as a Taylor series with finite number of initial terms leads to

$$\begin{aligned}
\eta(\xi)\ (\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi))) &= \eta(\xi) \sum_{l=0}^{k} \frac{1}{l!} \left(\frac{d^l}{dy^l} \left(\bar{\alpha}_k(x,y) + \beta_k(x,y)\right)\Big|_{(\xi,0)} \right) (\eta(\xi))^l \\
&= \sum_{l=0}^{k} \frac{1}{l!} \left(\frac{d^l}{dy^l} \left(\bar{\alpha}_k(x,y) + \beta_k(x,y)\right)\Big|_{(\xi,0)} \right) (\eta(\xi))^{l+1} \\
&= \sum_{l=0}^{k} \frac{1}{(l+1)!} \left(\frac{d^{l+1}}{dy^{l+1}} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \Big|_{(\xi,0)} \right) (\eta(\xi))^{l+1} \\
&= \sum_{l=1}^{k+1} \frac{1}{l!} \left(\frac{d^l}{dy^l} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \Big|_{(\xi,0)} \right) (\eta(\xi))^{l}
\end{aligned} \tag{4.45}$$

and thus to

$$\begin{aligned}
\left(\int_0^{h_i} ([\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_2)^2 \right)^{\frac{1}{2}} &\lesssim \left(\int_0^{h_i} \sum_{l=1}^{k+1} \left(\frac{d^l}{dy^l} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \Big|_{(\xi,0)} \right)^2 (\eta(\xi))^{2l} \right)^{\frac{1}{2}} \\
&\lesssim \left(\int_0^{h_i} \sum_{l=1}^{k+1} \left(\frac{d^l}{dy^l} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \Big|_{(\xi,0)} \right)^2 (h_i)^{2l(k+2)} \right)^{\frac{1}{2}} \\
&\lesssim \sum_{l=1}^{k+1} \| \frac{d^l}{dy^l} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \|_{0,E_i} \, (h_i)^{l(k+2)} \ .
\end{aligned}$$

Using the scaling argument (2.41) again, one obtains

$$\left(\int_0^{h_i} ([\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_2)^2 \right)^{\frac{1}{2}} \lesssim \sum_{l=1}^{k+1} \| \frac{d^l}{dy^l} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \|_{0,\mathbf{D}_i(\hat{T}_i)} \, (h_i)^{l(k+2)-\frac{1}{2}} \ . \tag{4.46}$$

For $q \in L^2(F_h^{-1}(\Gamma_i))$, it holds

$$\begin{aligned}
&\langle \mathbf{n} \cdot \hat{\mathbf{v}}_{h,i}, q \rangle_{0,F_h^{-1}(\Gamma_i)} = \langle \mathbf{n} \cdot \hat{\mathbf{w}}_{h,i}, q_{\mathbf{D}} \rangle_{0,\Gamma_{\mathbf{D},i}} \\
&= \int_0^{h_i} \left(-\eta'(\xi)(\alpha_k(\xi,\eta(\xi)) + \xi\, \beta_k(\xi,\eta(\xi))) + \eta(\xi)\ (\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)))\right) q_{\mathbf{D}}(\xi,\eta(\xi))\, d\xi \, .
\end{aligned}$$

where $q_{\mathbf{D}} = q \circ \mathbf{D}_i^{-1}$. Then, using the Cauchy-Schwarz inequality, (4.43) and (4.46) leads to

$$\begin{aligned}
\langle \mathbf{n} \cdot \hat{\mathbf{v}}_{h,i}, q \rangle_{0,F_h^{-1}(\Gamma_i)} &\leq \left[\left(\int_0^{h_i} (\eta'(\xi))^2 (\alpha_k(\xi,\eta(\xi)) + \xi\, \beta_k(\xi,\eta(\xi)))^2 d\xi \right)^{\frac{1}{2}} \right. \\
&\quad \left. + \left(\int_0^{h_i} (\eta(\xi))^2 (\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)))^2 d\xi \right)^{\frac{1}{2}} \right] \left(\int_0^{h_i} \ q_{\mathbf{D}}(\xi,\eta(\xi))^2 \, d\xi \right)^{\frac{1}{2}} \\
&\lesssim \left[h_i^{k+\frac{1}{2}} \| [\hat{\mathbf{w}}_{h,i}]_1 \|_{0,\mathbf{D}_i(\hat{T}_i)} + \sum_{l=1}^{k+1} \| \frac{d^l}{dy^l} [\hat{\mathbf{w}}_{h,i}(x,y)]_2 \|_{0,\mathbf{D}_i(\hat{T}_i)} \, (h_i)^{l(k+2)-\frac{1}{2}} \right] \| q_{\mathbf{D}} \|_{0,\Gamma_{\mathbf{D},i}} \, .
\end{aligned}$$

Using the following scaling argument

$$\begin{aligned}\|\frac{d^l}{dy^l}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\|_{0,\mathbf{D}_i(\hat{T}_i)} &\eqsim h_i^l\|\left(\frac{d^l}{dy^l}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\right)\circ\mathbf{F}_{\mathbf{D}_i(\hat{T}_i),\mathrm{ref}}\|_{0,\hat{T}_{\mathrm{ref}}}\\ &\eqsim h_i^l\|[\hat{\mathbf{w}}_{h,i}(x,y)]_2\circ\mathbf{F}_{\mathbf{D}_i(\hat{T}_i),\mathrm{ref}}\|_{0,\hat{T}_{\mathrm{ref}}}\\ &\eqsim h_i^l\|[\hat{\mathbf{w}}_{h,i}(x,y)]_2\|_{0,\mathbf{D}_i(\hat{T}_i)}\end{aligned} \tag{4.47}$$

leads to

$$\begin{aligned}\langle\mathbf{n}\cdot\hat{\mathbf{v}}_{h,i},q\rangle_{0,F_h^{-1}(\Gamma_i)} &\lesssim \left(h_i^{k+\frac{1}{2}}\|[\hat{\mathbf{w}}_{h,i}]_1\|_{0,\mathbf{D}_i(\hat{T}_i)}+\sum_{l=1}^{k+1}\|[\hat{\mathbf{w}}_{h,i}(x,y)]_2\|_{0,\mathbf{D}_i(\hat{T}_i)}\,(h_i)^{l(k+1)-\frac{1}{2}}\right)\|q_{\mathbf{D}}\|_{0,\Gamma_{\mathbf{D},i}}\\ &\lesssim \left(h_i^{k+\frac{1}{2}}\|[\hat{\mathbf{w}}_{h,i}]_1\|_{0,\mathbf{D}_i(\hat{T}_i)}+\sum_{l=1}^{k+1}\|[\hat{\mathbf{w}}_{h,i}(x,y)]_2\|_{0,\mathbf{D}_i(\hat{T}_i)}\,(h_i)^{k+\frac{1}{2}}\right)\|q\|_{0,F_h^{-1}(\Gamma_i)}\\ &\lesssim h_i^{k+\frac{1}{2}}\left(\|[\hat{\mathbf{w}}_{h,i}]_1\|_{0,\mathbf{D}_i(\hat{T}_i)}+\|[\hat{\mathbf{w}}_{h,i}(x,y)]_2\|_{0,\mathbf{D}_i(\hat{T}_i)}\right)\|q\|_{0,F_h^{-1}(\Gamma_i)}\\ &\lesssim h_i^{k+\frac{1}{2}}\|\hat{\mathbf{w}}_{h,i}\|_{0,\mathbf{D}_i(\hat{T}_i)}\|q\|_{0,F_h^{-1}(\Gamma_i)}\end{aligned}$$

and finally to the bound with respect to $L^2(F_h^{-1}(\Gamma_i))$:

$$\langle\mathbf{n}\cdot\hat{\mathbf{v}}_{h,i},q\rangle_{0,F_h^{-1}(\Gamma_i)}\lesssim h_i^{k+\frac{1}{2}}\|\hat{\mathbf{v}}_{h,i}\|_{0,\hat{T}_i}\|q\|_{0,F_h^{-1}(\Gamma_i)} \tag{4.48}$$

using using (4.37a). For a bound with respect to $H^1(F_h^{-1}(\Gamma_i))$, first note that integrating by parts leads to

$$\begin{aligned}&\int_0^{h_i}\eta'(\xi)[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1 q_{\mathbf{D}}(\xi,\eta(\xi))\,d\xi\\ &\quad=\int_0^{h_i}-\eta(\xi)[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\frac{d}{d\xi}q_{\mathbf{D}}(\xi,\eta(\xi))+[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\frac{d}{d\xi}\left(q_{\mathbf{D}}(\xi,\eta(\xi))\eta(\xi)\right)\,d\xi\\ &\quad=\int_0^{h_i}-\eta(\xi)[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\frac{d}{d\xi}q_{\mathbf{D}}(\xi,\eta(\xi))-q_{\mathbf{D}}(\xi,\eta(\xi))\eta(\xi)\frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\,d\xi\ .\end{aligned}$$

Thus, inserting this into (4.42b) leads to

$$\begin{aligned}\langle\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}_h,q\rangle_{0,F_h^{-1}(\Gamma_i)} &=\langle\mathbf{n}\cdot\hat{\mathbf{w}}_{h,i},q_{\mathbf{D}}\rangle_{0,\Gamma_{\mathbf{D},i}}\\ &=\int_0^{h_i}\left(-\eta'(\xi)[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1+[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_2\right)\,q_{\mathbf{D}}(\xi,\eta(\xi))\,d\xi\\ &=\int_0^{h_i}\eta(\xi)[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\frac{d}{d\xi}q_{\mathbf{D}}(\xi,\eta(\xi))+q_{\mathbf{D}}(\xi,\eta(\xi))\eta(\xi)\frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\,d\xi\\ &\quad+\int_0^{h_i}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_2\,q_{\mathbf{D}}(\xi,\eta(\xi))\,d\xi\\ &=\int_0^{h_i}\eta(\xi)[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\frac{d}{d\xi}q_{\mathbf{D}}(\xi,\eta(\xi))\,d\xi\\ &\quad+\int_0^{h_i}\eta(\xi)\left(\bar{\alpha}_k(\xi,\eta(\xi))+\beta_k(\xi,\eta(\xi))+\frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\right)\,q_{\mathbf{D}}(\xi,\eta(\xi))\,d\xi\,.\end{aligned}$$

Using the Cauchy Schwarz inequality, one obtains

$$\begin{aligned}\langle \hat{\mathbf{n}} \cdot \hat{\mathbf{v}}_h, q\rangle_{0,F_h^{-1}(\Gamma_i)} \leq & \left(\int_0^{h_i} \eta(\xi)^2 \left(\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)) + \frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\right)^2 d\xi\right)^{\frac{1}{2}} \|q\|_{0,F_h^{-1}(\Gamma_i)} \\ & + \left(\int_0^{h_i} \eta(\xi)^2([\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1)^2 d\xi\right)^{\frac{1}{2}} \left(\int_0^{h_i} \left(\frac{d}{d\xi} q(\xi,\eta(\xi))\right)^2 d\xi\right)^{\frac{1}{2}} \\ \lesssim & \, h_i^{k+2} \left(\int_0^{h_i} \left(\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)) + \frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\right)^2 d\xi\right)^{\frac{1}{2}} \|q\|_{0,F_h^{-1}(\Gamma_i)} \\ & + h_i^{k+2} \|[\hat{\mathbf{w}}_h]_1\|^2_{0,F_h^{-1}(\Gamma_i)} |q|_{1,F_h^{-1}(\Gamma_i)} \,. \end{aligned} \tag{4.49}$$

It remains to estimate the term in brackets. Therefore, first note that (4.45) implies

$$\begin{aligned}\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)) &= \sum_{l=1}^{k+1} \frac{1}{l!} \left(\frac{d^l}{dy^l}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)}\right)(\eta(\xi))^{l-1} \\ &= \sum_{l=0}^{k} \frac{1}{(l+1)!} \left(\frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)}\right)(\eta(\xi))^{l} \\ &= \frac{d}{dy}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)} + \sum_{l=1}^{k} \frac{1}{(l+1)!} \left(\frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)}\right)(\eta(\xi))^{l} \,. \end{aligned} \tag{4.50}$$

One the other hand, due to the fact that $y \mapsto \frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1$ is a polynom of order k, the taylor formula implies

$$\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1 = \sum_{k=0}^{k} \frac{1}{l!} \frac{\partial^l}{\partial y^l} \frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} y^l \tag{4.51}$$

such that

$$\begin{aligned}\frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1 =& \frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,\eta(\xi))} + \eta'(\xi)\frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,\eta(\xi))} \\ =& \sum_{l=0}^{k} \frac{1}{l!}\frac{\partial^l}{\partial y^l}\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} (\eta(\xi))^l + \eta'(\xi)\frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,\eta(\xi))} \\ =& \frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} + \sum_{l=1}^{k} \frac{1}{l!}\frac{\partial^l}{\partial y^l}\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} (\eta(\xi))^l \\ & + \eta'(\xi)\frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,\eta(\xi))} \,. \end{aligned}$$

Combining this with (4.50) leads to

$$
\begin{aligned}
&\int_0^{h_i} \left(\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)) + \frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\right)^2 d\xi \\
&\quad = \int_0^{h_i} \left(\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} + \sum_{l=1}^{k} \frac{1}{l!}\frac{\partial^l}{\partial y^l}\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} (\eta(\xi))^l + \frac{d}{dy}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)} \right. \\
&\qquad \left. + \eta'(\xi)\frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,\eta(\xi))} + \sum_{l=1}^{k} \frac{1}{(l+1)!}\left(\frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)}\right)(\eta(\xi))^l \right)^2 d\xi \\
&\quad = \int_0^{h_i} \left(\operatorname{div}\ \mathbf{w}_{h,i}(\xi,0) + \sum_{l=1}^{k} \frac{1}{l!}\frac{\partial^l}{\partial y^l}\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,0)} (\eta(\xi))^l \right. \\
&\qquad \left. + \eta'(\xi)\frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}(x,y)]_1\Big|_{(\xi,\eta(\xi))} + \sum_{l=1}^{k} \frac{1}{(l+1)!}\left(\frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}(x,y)]_2\Big|_{(\xi,0)}\right)(\eta(\xi))^l \right)^2 d\xi \\
&\quad \lesssim \|\operatorname{div}\hat{\mathbf{w}}_{h,i}\|^2_{0,E_i} + \sum_{l=1}^{k} h_i^{2l(k+2)} \left(\left\| \frac{\partial^l}{\partial y^l}\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}]_1 \right\|^2_{0,E_i} + \left\| \frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}]_2 \right\|^2_{0,E_i} \right) \\
&\qquad + h_i^{2(k+1)} \| \frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}(x,y)]_1 \|^2_{0,\Gamma_{\mathbf{D},i}} \ .
\end{aligned}
$$

Using again (2.41) and then (4.47), one obtains

$$
\begin{aligned}
&\int_0^{h_i} \left(\bar{\alpha}_k(\xi,\eta(\xi)) + \beta_k(\xi,\eta(\xi)) + \frac{d}{d\xi}[\hat{\mathbf{w}}_{h,i}(\xi,\eta(\xi))]_1\right)^2 d\xi \\
&\quad \lesssim h_i^{-1}\|\operatorname{div}\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} + \sum_{l=1}^{k} h_i^{2l(k+2)-1} \left(\left\| \frac{\partial^l}{\partial y^l}\frac{\partial}{\partial x}[\hat{\mathbf{w}}_{h,i}]_1 \right\|^2_{0,\mathbf{D}(\hat{T}_i)} + \left\| \frac{d^{l+1}}{dy^{l+1}}[\hat{\mathbf{w}}_{h,i}]_2 \right\|^2_{0,\mathbf{D}(\hat{T}_i)} \right) \\
&\qquad + h_i^{2k+1} \| \frac{\partial}{\partial y}[\hat{\mathbf{w}}_{h,i}]_1 \|^2_{0,\mathbf{D}(\hat{T}_i)} \\
&\quad \lesssim h_i^{-1}\|\operatorname{div}\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} + \sum_{l=1}^{k} h_i^{2l(k+1)-3} \left(\|[\hat{\mathbf{w}}_{h,i}]_1\|^2_{0,\mathbf{D}(\hat{T}_i)} + \|[\hat{\mathbf{w}}_{h,i}]_2\|^2_{0,\mathbf{D}(\hat{T}_i)} \right) \\
&\qquad + h_i^{2k-1} \|[\hat{\mathbf{w}}_{h,i}]_1\|^2_{0,\mathbf{D}(\hat{T}_i)} \\
&\quad \lesssim h_i^{-1}\|\operatorname{div}\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} + \sum_{l=1}^{k} h_i^{2l(k+1)-3} \|\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} + + h_i^{2k-1} \|\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} \\
&\quad \lesssim h_i^{-1}\|\operatorname{div}\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} + h_i^{2k-1} \|\hat{\mathbf{w}}_{h,i}\|^2_{0,\mathbf{D}(\hat{T}_i)} \\
&\quad \lesssim h_i^{-1}\|\operatorname{div}\hat{\mathbf{v}}_{h,i}\|^2_{0,\hat{T}_i} + h_i^{2k-1} \|\hat{\mathbf{v}}_{h,i}\|^2_{0,\hat{T}_i} \ .
\end{aligned}
$$

Inserting this in (4.49) finally leads to

$$\begin{aligned}\langle \hat{\mathbf{n}} \cdot \hat{\mathbf{v}}_h, q\rangle_{0,F_h^{-1}(\Gamma_i)} \lesssim\, & h_i^{k+2} \left(h_i^{-1} \|\operatorname{div} \hat{\mathbf{v}}_{h,i}\|_{0,\hat{T}_i}^2 + h_i^{2k-1} \|\hat{\mathbf{v}}_{h,i}\|_{0,\hat{T}_i}^2 \right)^{\frac{1}{2}} \|q\|_{0,F_h^{-1}(\Gamma_i)} \\ & + h_i^{k+2} \|[\hat{\mathbf{w}}_h]_1\|_{0,F_h^{-1}(\Gamma_i)}^2 |q|_{1,F_h^{-1}(\Gamma_i)} \\ \lesssim\, & h_i^{k+\frac{3}{2}} \left(\left(\|\operatorname{div} \hat{\mathbf{v}}_h\|_{0,\hat{T}_i} + \|\hat{\mathbf{v}}_h\|_{0,\hat{T}_i} \right) \|q\|_{0,F_h^{-1}(\Gamma_i)} + \|\hat{\mathbf{v}}_h\|_{0,\hat{T}_i} |q|_{1,F_h^{-1}(\Gamma_i)} \right) \\ \lesssim\, & h_i^{k+\frac{3}{2}} \left(\|\operatorname{div} \hat{\mathbf{v}}_h\|_{0,\hat{T}_i} + \|\hat{\mathbf{v}}_h\|_{0,\hat{T}_i} \right) \left(\|q\|_{0,F_h^{-1}(\Gamma_i)} + |q|_{1,F_h^{-1}(\Gamma_i)} \right) \\ \lesssim\, & h_i^{k+\frac{3}{2}} \left((\|\operatorname{div} \hat{\mathbf{v}}_h\|_{0,\hat{T}_i} + \|\hat{\mathbf{v}}_h\|_{0,\hat{T}_i})^2 \right)^{\frac{1}{2}} \left((\|q\|_{0,F_h^{-1}(\Gamma_i)} + |q|_{1,F_h^{-1}(\Gamma_i)})^2 \right)^{\frac{1}{2}} \\ \lesssim\, & h_i^{k+\frac{3}{2}} \left(\|\operatorname{div} \hat{\mathbf{v}}_h\|_{0,\hat{T}_i}^2 + \|\hat{\mathbf{v}}_h\|_{0,\hat{T}_i}^2 \right)^{\frac{1}{2}} \left(\|q\|_{0,F_h^{-1}(\Gamma_i)}^2 + |q|_{1,F_h^{-1}(\Gamma_i)}^2 \right)^{\frac{1}{2}} . \end{aligned} \tag{4.52}$$

The final step consists in transfering the inequalities (4.48) and (4.52) back to Γ_T. Therfore, note that due to the construction of $\mathbf{v}_{h,i}$ (i.e. (4.36)), it holds

$$\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma_i} = \langle \hat{\mathbf{n}} \cdot \hat{\mathbf{v}}_h, q\rangle_{0,F_h^{-1}(\Gamma_i)} \tag{4.53}$$

and combining Theorem 1.8 and (4.10) implies

$$\|\hat{\mathbf{v}}_h\|_{0,\hat{T}_i} \lesssim \|\mathbf{v}_h\|_{0,T_i} , \tag{4.54a}$$

$$\|\hat{\operatorname{div}}\, \hat{\mathbf{v}}_h\|_{0,\hat{T}_i} \lesssim \|\operatorname{div} \mathbf{v}_h\|_{0,T_i} . \tag{4.54b}$$

Thus, (4.48) turns into

$$\langle \mathbf{n} \cdot \mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} \lesssim h_i^{k+\frac{1}{2}} \|\mathbf{v}_{h,i}\|_{0,T_i} \|\tilde{q}\|_{0,\Gamma_i} \quad \forall \tilde{q} \in L^2(\Gamma_i) \tag{4.55}$$

and (4.52) turns into

$$\langle \mathbf{n} \cdot \mathbf{v}_{h,i}, q\rangle_{0,\Gamma_i} \lesssim h_i^{k+\frac{3}{2}} \left(\|\operatorname{div} \mathbf{v}_h\|_{0,T_i}^2 + \|\mathbf{v}_h\|_{0,T_i}^2 \right)^{\frac{1}{2}} \|\tilde{q}\|_{1,\Gamma_i} \quad \forall \tilde{q} \in H^1(\Gamma_i) \tag{4.56}$$

Replacing T_i by $\widetilde{T_i}$ since $\mathbf{v}_{h,i}$ belong to a finite-dimensional space and summing over all boundary edges leads to

$$\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma} \lesssim h^{k+\frac{1}{2}} \|\mathbf{v}_h\|_{0,\Omega} \|q\|_{0,\Gamma} \quad \forall q \in L^2(\Gamma) \tag{4.57}$$

and

$$\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma} \lesssim h_i^{k+\frac{3}{2}} \left(\|\operatorname{div} \mathbf{v}_h\|_{0,\Omega}^2 + \|\mathbf{v}_h\|_{0,\Omega}^2 \right)^{\frac{1}{2}} \|q\|_{1,\Gamma} \quad \forall q \in H^1(\Gamma). \tag{4.58}$$

Finally, using the fact that $H^{\frac{1}{2}}(\Gamma)$ is an interpolation space of type $\frac{1}{2}$ for $L^2(\Gamma)$ and $H^1(\Gamma)$ as mentioned in Theorem 1.3 and applying it as in the proof of Theorem 2.1 leads to

$$|\langle \mathbf{n} \cdot \mathbf{v}_h, q\rangle_{0,\Gamma}| \lesssim h^{k+1} \left(\|\mathbf{v}_h\|_{0,\Omega}^2 + \|\operatorname{div} \mathbf{v}_h\|_{0,\Omega}^2 \right)^{\frac{1}{2}} \|q\|_{\frac{1}{2},\Gamma}. \tag{4.59}$$

□

Chapter 5

Least Squares Method with Parametric Finite Elements

The aim of this chapter is to show how the (iso)parametric elements derived in the previous chapter can be used for the error analysis of the first-order system least squares finite-element approximation. The main result states that the use of the polynomial approximation is sufficient to retain the optimal order of convergence for these elements.

The first step therefore is to construct an implementable approximation of the least squares functional $\mathcal{F}$, and this will be done in the first section. In particular the $L^2(\Omega_h)$-orthogonal projection onto the space of discontinuous isoparametric piecewise polynomials of degree k on $\mathcal{T}_h$ is used to approximate the right hand side f.

In the second section, the main theorem of this chapter is then proven. Therefore, a lower bound for the least squares functional is derived using the estimate for the normal flux of the finite-element function on the approximated boundary from Theorem 4.3. Further, a connection between the least squares functional and its approximation is established and finally an upper bound for the approximated functional is stated, using the interpolation properties of the finite-element interpolation operators derived in Chapter 4.

This theoretical result is then illustrated by a numerical example in the third section.

5.1 Approximation of the Least Squares Functional

Similarly to the approach in Chapter 3, the next step after the construction of Ω_h and the definition of the corresponding (iso)parametric elements in order to apply the least squares method on domains with curved boundaries is to define an implementable approximated functional $\mathcal{F}_h$. Therefore the integration domain Ω is replaced by Ω_h and the right hand side has to be replaced by an implementable one, such that

$$\mathcal{F}_h(\mathbf{u}, p) = \|\operatorname{div} \mathbf{u} - f_h\|^2_{0,\Omega_h} + \|\mathbf{u} + \nabla p\|^2_{0,\Omega_h} \,. \tag{5.1}$$

has to be minimized with respect of $(\mathbf{u}_h, p_h) \in \mathbf{V}_h^k \times \dot{Q}_h^{k+1}$. The $L^2(\Omega)$-orthogonal projection approximate f with the desired approximation order but it is not computable without further assumptions, contrary to the $L^2(\Omega_h)$-orthogonal projection onto the space of discontinuous isoparametric piecewise polynomials of degree k on $\mathcal{T}_h$. Figure 5.1 illustrates the difference between the $L^2(\Omega_h)$ and the $L^2(\Omega)$-orthogonal projection for $k = 1$. The following lemma (see [8]) states that choosing f_h as the $L^2(\Omega_h)$-orthogonal projection onto the space of discontinuous

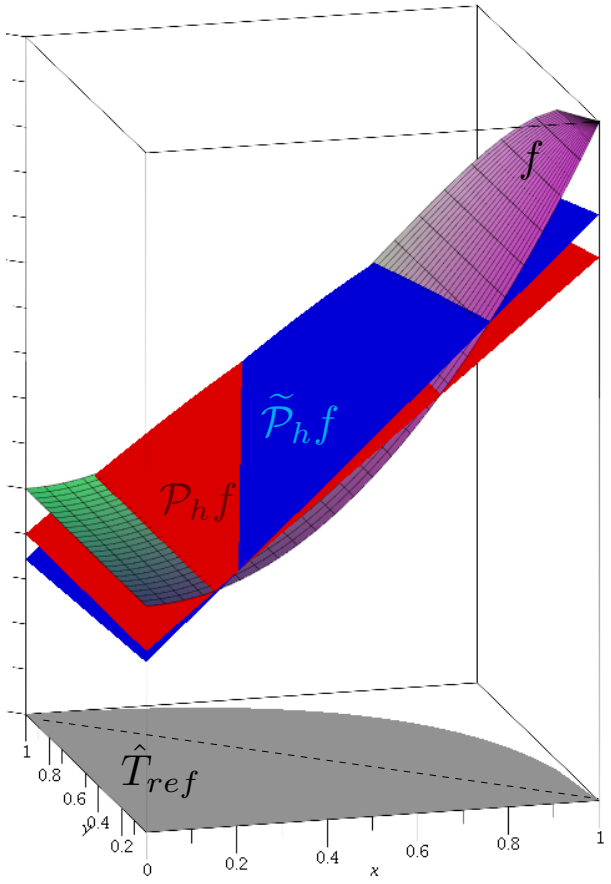


Figure 5.1: Difference between $L^2(\Omega_h)$ and $L^2(\Omega)$-orthogonal projection

isoparametric piecewise polynomials of degree k on $\mathcal{T}_h$ leads to a sufficiently good approximation to f on Ω.

Lemma 5.1. *Let $f \in W^1_\infty(\Omega \cup \Omega_h) \cap H^{k+1}(\Omega \cup \Omega_h)$, where Ω_h is the polynomial approximation of Ω of degree $k+1$ defined in the previous chapter with the corresponding triangulation $\mathcal{T}_h$. If $\mathcal{P}_h : L^2(\Omega_h) \to Z_h^k$ denotes the $L^2(\Omega_h)$-orthogonal projection onto the space of (discontinuous) isoparametric piecewise polynomials of degree k on $\mathcal{T}_h$, then it holds*

$$\|f - \mathcal{P}_h f\|_{0,\Omega} \lesssim h^{k+1} \left(\|f\|_{W^1_\infty(\Omega \cup \Omega_h)} + \|f\|_{k+1,\Omega \cup \Omega_h} \right) . \tag{5.2}$$

Proof: Consider $T_i \in \mathcal{T}_h$ and the corresponding curved triangle $\widetilde{T}_i$. Then, first note that

$$\begin{aligned} \|f - \mathcal{P}_h f\|_{0,\widetilde{T}_i} &\lesssim \|(f - \mathcal{P}_h f) \circ \mathbf{\Phi}_h\|_{0,T_i} \\ &\leq \|f \circ \mathbf{\Phi}_h - f\|_{0,T_i} + \|f - \mathcal{P}_h f\|_{0,T_i} + \|\mathcal{P}_h f - (\mathcal{P}_h f) \circ \mathbf{\Phi}_h\|_{0,T_i} \end{aligned} \tag{5.3}$$

holds, with the mapping $\mathbf{\Phi}_h$ defined in Chapter 4. Now, these terms are estimated separately. For the first term, using

$$\begin{aligned} (f \circ \mathbf{\Phi}_h - f)(\mathbf{x}) &= \int_0^1 \frac{d}{ds} f\left(\mathbf{x} + s(\mathbf{\Phi}_h(\mathbf{x}) - \mathbf{x})\right) \, ds \\ &= \int_0^1 \nabla f\left(\mathbf{x} + s(\mathbf{\Phi}_h(\mathbf{x}) - \mathbf{x})\right) \, ds \cdot (\mathbf{\Phi}_h(\mathbf{x}) - \mathbf{x}) \, , \end{aligned} \tag{5.4}$$

one obtains

$$\|f \circ \mathbf{\Phi}_h - f\|_{0,T_i} \leq \|\nabla f\|_{L^\infty(T_i \cup \tilde{T}_i)} \|\mathbf{\Phi}_h - \mathrm{id}\|_{0,T_i} \ . \tag{5.5}$$

Using (4.12a) leads to

$$\|f \circ \mathbf{\Phi}_h - f\|_{0,T_i} \lesssim h^{k+1} \|f\|_{W^1_\infty(T_i \cup \tilde{T}_i)} \ . \tag{5.6}$$

The second term can be bounded using the best approximation property of the orthogonal projection:

$$\|f - \mathcal{P}_h f\|_{0,T_i} = \inf_{z_h \in Z_h^k} \|f - z_h\|_{0,T_i} \lesssim h^{k+1} \|f\|_{k+1,T_i} \ , \tag{5.7}$$

and for the third term, using the same way as in (5.4), one obtains

$$\begin{aligned} \|\mathcal{P}_h f - (\mathcal{P}_h f) \circ \mathbf{\Phi}_h\|_{0,T_i} &\lesssim \left\| \int_0^1 \frac{d}{ds} \mathcal{P}_h \left(\mathbf{x} + s(\mathbf{\Phi}_h(\mathbf{x}) - \mathbf{x}) \right) \, ds \right\|_{0,T_i} \\ &\lesssim \left\| \int_0^1 \nabla \mathcal{P}_h \left(\mathbf{x} + s(\mathbf{\Phi}_h(\mathbf{x}) - \mathbf{x}) \right) \, ds \cdot (\mathbf{\Phi}_h(\mathbf{x}) - \mathbf{x}) \right\|_{0,T_i} \\ &\lesssim \|\nabla \mathcal{P}_h\|_{L^\infty(T_i \cup \tilde{T}_i)} \|\mathbf{\Phi}_h - \mathrm{id}\|_{0,T_i} \end{aligned}$$

Using (4.12a) again leads to

$$\|\mathcal{P}_h f - (\mathcal{P}_h f) \circ \mathbf{\Phi}_h\|_{0,T_i} \lesssim h^{k+1} \|\mathcal{P}_h f\|_{W^1_\infty(T \cup \tilde{T}_i)} \ . \tag{5.8}$$

Inserting (5.6), (5.7) and (5.8) into (5.3) leads to

$$\|f - \mathcal{P}_h f\|_{0,\widetilde{T}_i} \lesssim h^{k+1} \left(\|f\|_{W^1_\infty(T_i \cup \tilde{T}_i)} + \|f\|_{k+1,T_i} + \|\mathcal{P}_h f\|_{W^1_\infty(T \cup \tilde{T}_i)} \right) \ . \tag{5.9}$$

Further, recall that due to the properties of the orthogonal projection it holds

$$\|\mathcal{P}_h f\|_{L^2(T_{\mathrm{ref}})} \leq \|f\|_{L^2(T_{\mathrm{ref}})} \ . \tag{5.10}$$

Therefore, using the fact on a finite-dimensional space that all norms are equivalent, one obtains

$$\begin{aligned} \|\mathcal{P}_h f\|_{W^1_\infty(T_i \cup \tilde{T}_i)} &\lesssim \|\mathcal{P}_h f\|_{W^1_\infty(T_{\mathrm{ref}} \cup \tilde{T}_{\mathrm{ref}})} \lesssim \|f\|_{L^2(T_{\mathrm{ref}})} \lesssim \|f\|_{W^1_\infty(T_{\mathrm{ref}})} \\ &\lesssim \|f\|_{W^1_\infty(T_{\mathrm{ref}} \cup \tilde{T}_{\mathrm{ref}})} \lesssim \|f\|_{W^1_\infty(T_i \cup \tilde{T}_i)} \ . \end{aligned}$$

Inserting this into (5.9) leads to

$$\|f - \mathcal{P}_h f\|_{0,\tilde{T}_i} \lesssim h^{k+1} \left(\|f\|_{W^1_\infty(T_i \cup \tilde{T}_i)} + \|f\|_{k+1,T_i} \right) \ . \tag{5.11}$$

Summing over all triangles concludes the proof. □

5.2 Error Analysis of the LSFEM with Parametric Finite Elements

The aim of this section is to derive a theorem similar to Theorem 1.9 for the domain Ω with curved boundaries introduced in Chapter 4. Similarly to Chapter 3, a lower bound of the least squares functional has to be derived, as the bound derived in Section 1.5 does not hold due to the fact that the normal flux does not vanish on Γ. However, a lower bound can be given using the estimate from the Theorem 4.3, as the following lemma (see [8]) states:

Lemma 5.2. *Let* $(\mathbf{u},p)\in H_\Gamma(\mathrm{div},\Omega)\times\dot{H}^1(\Omega)$ *denote the exact solution of the system (1.33). Then, it holds*

$$\mathcal{F}(\mathbf{v}_h,q_h)+h^{2k+2}\left(\|\mathbf{u}\|^2_{0,\Omega}+\|\mathrm{div}\,\mathbf{u}\|^2_{0,\Omega}\right)\gtrsim\|\mathbf{v}_h-\mathbf{u}\|^2_{\mathrm{div},\Omega}+\|\nabla(q_h-p)\|^2_{0,\Omega} \tag{5.12}$$

for all $(\mathbf{v}_h,q_h)\in\mathbf{V}_h^k\times\dot{Q}_h^{k+1}$.

Proof: First note that due to the fact that the exact solution solves (1.33), it can be inserted in $\mathcal{F}(\mathbf{v}_h,q_h)$:

$$\begin{aligned}\mathcal{F}(\mathbf{u}_h,p_h)&=\|\mathrm{div}\,\mathbf{u}_h-f\|^2_{0,\Omega}+\|\mathbf{u}_h+\nabla p_h\|^2_{0,\Omega}\\&=\|\mathrm{div}(\mathbf{u}_h-\mathbf{u})\|^2_{0,\Omega}+\|\mathbf{u}_h-\mathbf{u}+\nabla(p_h-p)\|^2_{0,\Omega}\end{aligned}$$

The same idea as in Section 1.5 is used to get the lower bound, i.e., the term $\|\mathbf{v}_h-\mathbf{u}+\nabla(q_h-p)\|^2_{0,\Omega}$ has to be expanded and the mixed terms that then appear have to be distributed on the other terms, and in particular weighted against $\|\mathrm{div}(\mathbf{v}_h-\mathbf{u})\|^2_{0,\Omega}$. Here, the same weights are used as in Section 1.5 such that it holds (see (3.19))

$$\begin{aligned}\mathcal{F}(\mathbf{v}_h,q_h)\geq\frac{1}{2}\|\mathrm{div}(\mathbf{v}_h-\mathbf{u})\|^2_{0,\Omega}+\alpha\|\mathbf{v}_h-\mathbf{u}\|^2_{0,\Omega}\\+\alpha(1-2\alpha C_\Omega)\|\nabla(q_h-p)\|^2_{0,\Omega}-2\alpha\left|\langle\mathbf{n}\cdot\mathbf{v}_h,q_h-p\rangle_{0,\Gamma}\right|\ .\end{aligned} \tag{5.13}$$

with any $\alpha\in[0,1]$. Using Theorem 4.3 leads to

$$2\left|\langle\mathbf{n}\cdot\mathbf{v}_h,q_h-p\rangle_{0,\Gamma}\right|\leq 2C_S h^{k+1}\left(\|\mathbf{v}_h\|^2_{0,\Omega}+\|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right)^{1/2}\|q_h-p\|_{1/2,\Gamma}\ .$$

Again, using the same weights as in Lemma 3.2 one obtains

$$2\left|\langle\mathbf{n}\cdot\mathbf{v}_h,q_h-p\rangle_{0,\Gamma}\right|\leq Ch^2\left(\|\mathbf{v}_h\|^2_{0,\Omega}+\|\mathrm{div}\,\mathbf{v}_h\|^2_{0,\Omega}\right)+\frac{1}{2}\|\nabla(q_h-p)\|^2_{0,\Omega}\,.$$

with C=$8C_T^2C_\Omega^2C_S^2$. Inserting this in (5.13) and using the same further steps as in Lemma 3.2 gives the result. □

On the other hand, an upper bound has to be found for the least squares functional

$$\mathcal{F}_h(\mathbf{u}_h,p_h)=\|\mathrm{div}\,\mathbf{u}_h-f_h\|^2_{0,\Omega_h}+\|\mathbf{u}_h+\nabla p_h\|^2_{0,\Omega_h}\,, \tag{5.14}$$

where f_h is the $L^2(\Omega_h)$-orthogonal projection of f onto the space of (discontinuous) isoparametric piecewise polynomials of degree k on $\mathcal{T}_h$. This is the statement of the following lemma (see [8]):

Lemma 5.3. *Let* $(\mathbf{u},p)\in H_\Gamma(\mathrm{div},\Omega)\times\dot{H}^1(\Omega)$ *denote the exact solution of the system (1.33) and assume that it satisfies* $p\in H^2(\Omega)$. *Further, let* $(\mathbf{u}_h,p_h)$ *be the minimizer of* $\mathcal{F}_h(\mathbf{v}_h,q_h)$ *under all* $(\mathbf{v}_h,p_h)\in\mathbf{V}_h^k\times\dot{Q}_h^{k+1}$. *Then, it holds*

$$\mathcal{F}_h(\mathbf{u}_h,p_h)\lesssim h^{2k+2}\left(|\mathrm{div}\ \mathbf{u}|^2_{k+1,\Omega}+\|\mathbf{u}\|^2_{k+1,\Omega}+\|p\|^2_{k+2,\Omega}\right)\ . \tag{5.15}$$

Proof: Similarly to the proof of Lemma 3.3, the idea is to use the fact that the solution $(\mathbf{u}_h,p_h)$ minimize the approximated least squares functional and thus to bound the least squares

functional with the interpolation of the exact solution. To this end, first map the exact solution on Ω_h, using the mapping $\mathbf{\Phi}_h : \Omega \to \Omega_h$ defined in (4.11):

$$\widetilde{\mathbf{u}} = \left(\frac{1}{\det J_{\mathbf{\Psi}_h}} J_{\mathbf{\Psi}_h} \mathbf{u} \right) \circ \mathbf{\Phi}_h \,, \tag{5.16}$$

$$\widetilde{p} = p \circ \mathbf{\Phi}_h \,. \tag{5.17}$$

Due to Theorem 1.8,

$$\operatorname{div} \widetilde{\mathbf{u}} = \left(\frac{1}{\det J_{\mathbf{\Psi}_h}} \operatorname{div} \mathbf{u} \right) \circ \mathbf{\Phi}_h \,, \tag{5.18}$$

$$\nabla \widetilde{p} = \left(J^T_{\mathbf{\Phi}_h} \nabla p \right) \circ \mathbf{\Phi}_h \,. \tag{5.19}$$

Now, using the interpolation operators from Section 4.2,

$$\begin{aligned}
\mathcal{F}_h(\mathbf{u}_h, p_h) \leq & \mathcal{F}_h(\mathcal{R}_h \widetilde{\mathbf{u}}, \mathcal{I}_h \widetilde{p}) \\
= & \|\operatorname{div}(\mathcal{R}_h \widetilde{\mathbf{u}}) - f_h\|^2_{0,\Omega_h} + \|\mathcal{R}_h \widetilde{\mathbf{u}} + \nabla(\mathcal{I}_h \widetilde{p})\|^2_{0,\Omega_h} \\
= & \|\operatorname{div}(\mathcal{R}_h \widetilde{\mathbf{u}}) - \operatorname{div} \tilde{\mathbf{u}} + \operatorname{div} \tilde{\mathbf{u}} - f \circ \mathbf{\Phi}_h + f \circ \mathbf{\Phi}_h - f_h\|^2_{0,\Omega_h} \\
& + \|\mathcal{R}_h \widetilde{\mathbf{u}} - \tilde{\mathbf{u}} + \tilde{\mathbf{u}} - \nabla \tilde{p} + \nabla \tilde{p} \nabla(\mathcal{I}_h \widetilde{p})\|^2_{0,\Omega_h} \\
\lesssim & \|\operatorname{div}(\mathcal{R}_h \widetilde{\mathbf{u}} - \tilde{\mathbf{u}})\|^2_{0,\Omega_h} + \|\frac{1}{\det J_{\mathbf{\Psi}_h}} \operatorname{div} \mathbf{u} - f\|^2_{0,\Omega} + \|f \circ \mathbf{\Phi}_h - f_h\|^2_{0,\Omega_h} \\
& + \|\mathcal{R}_h \widetilde{\mathbf{u}} - \tilde{\mathbf{u}}\|^2_{0,\Omega_h} + \|\frac{1}{\det J_{\mathbf{\Psi}_h}} J_{\mathbf{\Psi}_h} \mathbf{u} - J^T_{\mathbf{\Phi}_h} \nabla p\|^2_{0,\Omega} + \|\nabla(\widetilde{p} - \mathcal{I}_h \widetilde{p})\|^2_{0,\Omega_h} \,,
\end{aligned} \tag{5.20}$$

where the transformed derivatives from (5.19) are inserted. For the second term, using (4.13b) one obtains

$$\|\frac{1}{\det J_{\mathbf{\Psi}_h}} \operatorname{div} \mathbf{u} - f\|^2_{0,\Omega} = \|\frac{1}{\det J_{\mathbf{\Psi}_h}} f - f\|^2_{0,\Omega} \lesssim h^{2k+2} \|f\|^2_{0,\Omega} \,. \tag{5.21}$$

For the third term, first note that

$$\|f_h \circ \mathbf{\Psi}_h - f\|^2_{0,\Omega} \leq \|f - f_h\|^2_{0,\Omega} + \|f_h - f_h \circ \mathbf{\Psi}_h\|^2_{0,\Omega} \tag{5.22}$$

holds, and that the first term can be estimated with Lemma 5.1. Due to

$$f_h - f_h \circ \mathbf{\Psi}_h = \int_0^1 \nabla f_h(x + \alpha(\mathbf{\Psi}_h(\mathbf{x}) - x))^T (\mathbf{\Psi}_h(\mathbf{x}) - x) \, d\alpha$$

this leads to

$$\|f_h \circ \mathbf{\Psi}_h - f\|^2_{0,\Omega} \lesssim h^{2k+2} \left(\|f\|^2_{W^1_\infty(\Omega \cup \Omega_h)} + \|f\|^2_{k+1,\Omega \cup \Omega_h} \right) \,. \tag{5.23}$$

For the fifth term, note that

$$\|\frac{1}{\det J_{\mathbf{\Psi}_h}} J_{\mathbf{\Psi}_h} \mathbf{u} - J^T_{\mathbf{\Phi}_h} \nabla p\|^2_{0,\Omega} = \|\frac{1}{\det J_{\mathbf{\Psi}_h}} J_{\mathbf{\Psi}_h} \mathbf{u} - J^T_{\mathbf{\Phi}_h} \mathbf{u}\|^2_{0,\Omega} \tag{5.24}$$

holds, due to the fact that $\mathbf{u}$ is the exact solution. This turns into

$$\|\frac{1}{\det J_{\mathbf{\Psi}_h}} J_{\mathbf{\Psi}_h} \mathbf{u} - J^T_{\mathbf{\Phi}_h} \nabla p\|^2_{0,\Omega} = \|\frac{1}{\det J_{\mathbf{\Psi}_h}} J_{\mathbf{\Psi}_h} - J^T_{\mathbf{\Phi}_h}\|^2_{L^\infty(\Omega)} \|\mathbf{u}\|^2_{0,\Omega} \lesssim h^{2k+2} \|\mathbf{u}\|^2_{0,\Omega} \tag{5.25}$$

using (4.13b) again. Estimating the first, fourth and sixth of the terms in (5.20) by the interpolation estimates from (4.25) and (4.27) and inserting (5.21), (5.23) and (5.25) leads to the result. □

The next theorem (main result in [8]) states that the optimal order of convergence of the least squares method described in Chapter 1 is retained for domains with curved boundaries by using the polynomial approximation of the domain and the corresponding (iso)parametric elements.

Theorem 5.1. *Let* $(\mathbf{u}, p) \in H_\Gamma(\operatorname{div}, \Omega) \times \dot{H}^1(\Omega)$ *denote the exact solution of the system (1.33) and assume that it satisfies* $p \in H^{k+2}(\Omega)$ *and* $\operatorname{div} \mathbf{u} \in H^{k+1}(\Omega)$. *Moreover, assume that* $f \in W^1_\infty(\Omega \cup \Omega_h)$ *and let* $(\mathbf{u}_h, p_h) \in \mathbf{V}_h^k \times \dot{Q}_h^{k+1}$ *be the (iso)parametric finite-element approximation minimizing* $\mathcal{F}_h(\mathbf{u}_h, p_h)$. *Then, it holds*

$$\begin{aligned}\|\mathbf{u} - \mathbf{u}_h\|_{\operatorname{div},\Omega} + \|p - p_h\|_{1,\Omega} \lesssim h^{k+1} \big(&\|\mathbf{u}\|_{k+1,\Omega} + |\operatorname{div} \mathbf{u}|_{k+1,\Omega} + \|p\|_{k+2,\Omega} \\ &+ \|f\|_{W^1_\infty(\Omega\cup\Omega_h)} + \|f\|_{k+1,\Omega\cup\Omega_h}\big) \ .\end{aligned} \tag{5.26}$$

Proof: Similarly to the proof of Theorem 3.1, the least squares functional and its approximation need to be connected in order to use the two previous lemmas. This can be done by

$$\begin{aligned}\mathcal{F}(\mathbf{u}_h, p_h) &= \|\operatorname{div} \mathbf{u}_h - f\|^2_{0,\Omega} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\Omega} \\ &\le \|f - f_h\|^2_{0,\Omega} + \|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,\Omega} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\Omega} \\ &= \|f - f_h\|^2_{0,\Omega} + \sum_{i=1}^{\bar{N}} \Big(\|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,\widetilde{T}_i} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\widetilde{T}_i}\Big) \\ &\lesssim \|f - f_h\|^2_{0,\Omega} + \sum_{i=1}^{\bar{N}} \Big(\|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,T_i} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,T_i}\Big) \\ &= \|f - f_h\|^2_{0,\Omega} + \|\operatorname{div} \mathbf{u}_h - f_h\|^2_{0,\Omega_h} + \|\mathbf{u}_h + \nabla p_h\|^2_{0,\Omega_h} \\ &= \|f - f_h\|^2_{0,\Omega} + \mathcal{F}_h(\mathbf{u}_h, p_h) \\ &\lesssim h^{2k+2} \left(\|f\|^2_{W^1_\infty(\Omega\cup\Omega_h)} + \|f\|^2_{k+1,\Omega\cup\Omega_h}\right) + \mathcal{F}_h(\mathbf{u}_h, p_h) \ ,\end{aligned} \tag{5.27}$$

where lemma 5.1 and the equivalence of the $L^2(\widetilde{T}_i)$ and $L^2(T_i)$ norms on finite dimensional spaces are used. Now, using (5.15) gives

$$\begin{aligned}\mathcal{F}(\mathbf{u}_h, p_h) \lesssim h^{2k+2} \big(&|\operatorname{div} \mathbf{u}|^2_{k+1,\Omega} + \|\mathbf{u}\|^2_{k+1,\Omega} + \|p\|^2_{k+2,\Omega} \\ &+ \|f\|_{W^1_\infty(\Omega\cup\Omega_h)} + \|f\|_{k+1,\Omega\cup\Omega_h)}\big) \ .\end{aligned} \tag{5.28}$$

Combining this with Lemma 5.2 in the same way as in Theorem 3.1 finishes the proof. □

5.3 Computational Results

In this section, Theorem 5.1 is confirmed numerically. To this end, consider the same problem as in Section 3.3, i.e. the Poisson problem on the unit disk with boundary conditions $\mathbf{n} \cdot \mathbf{u} = 0$ on $\partial\Omega$. The right-hand side is chosen to be $f(x_1, x_2) = \sin x_1$.

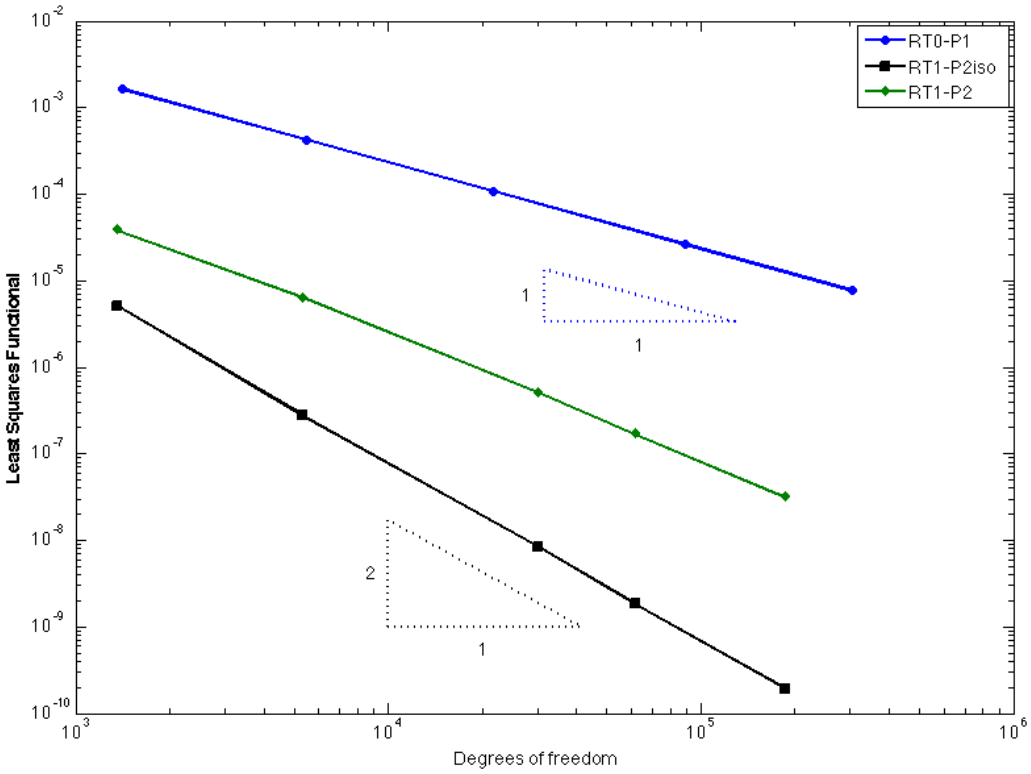


Figure 5.2: Reduction of least squares functional $\mathcal{F}_h(\mathbf{u}_h, p_h)$ for $f(x_1, x_2) = \sin\ x_1$

Figure 5.2 recalls the optimal convergence of the least squares functional for the lowest-order case and the suboptimal one for the higher-order case with the polygonal approximation of Ω. The lower curve shows the optimal order of convergence for the parametric RT_1 elements combined with continuous quadratic isoparametric elements that confirms Theorem 5.1. A further test is on a flower-shape domain

$$\Omega = \left\{ \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \ : \ -\left(6 + \frac{1}{2}\cos\left(12\ \tan^{-1}\left(\frac{x_2}{x_1}\right)\right)\right) + \|\mathbf{x}\| \right\} \ . \tag{5.29}$$

Figure 5.3 shows the reduction of the Least Squares Functional for $f(x_1, x_2) = x_1$ in the lower-order case and in the next higher order case (parametrics RT_1 and isoparametrics P_2). This confirms that the optimal convergence rates are retained in both cases. The solution on the coarsest mesh is presented in Figure 5.4.

.

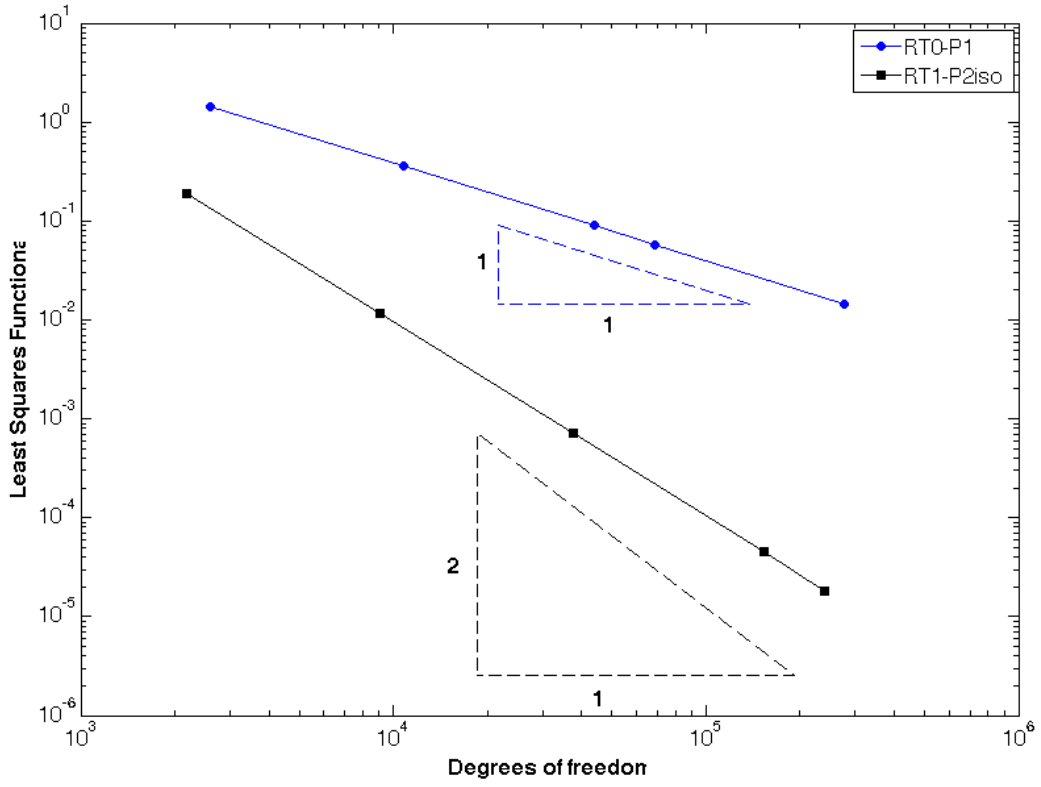
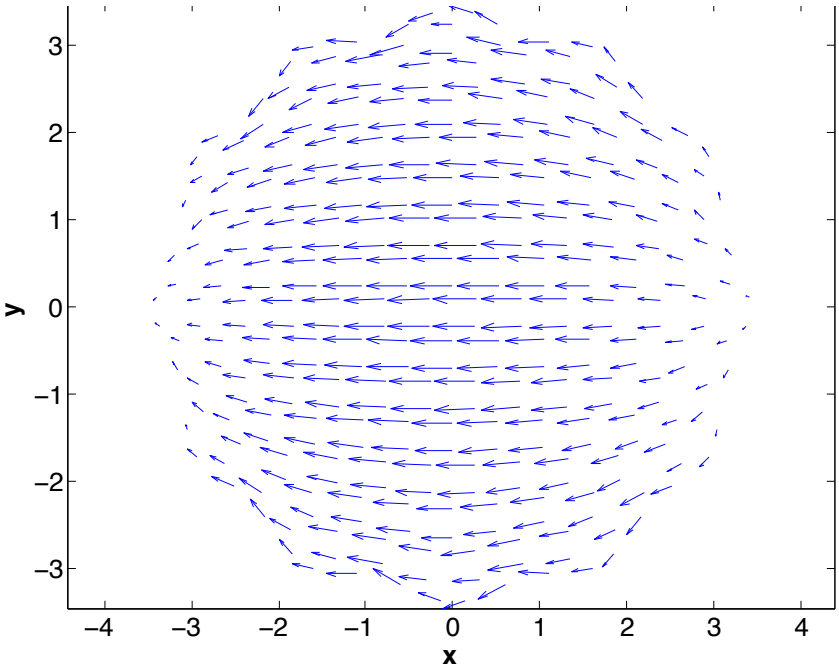
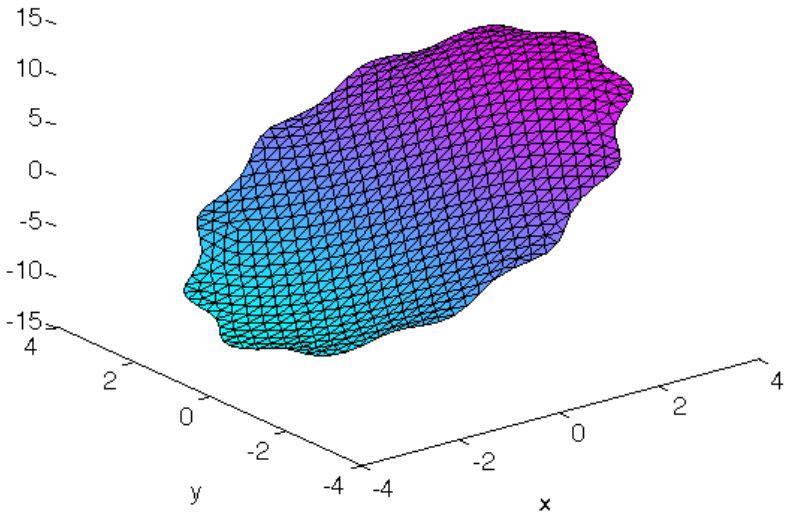


Figure 5.3: Reduction of least squares functional $\mathcal{F}_h(\mathbf{u}_h, p_h)$ on the flower-shape domain



(a) **u**



(b) p

Figure 5.4: Minimizer of (5.1) on the flower-shape domain

Chapter 6

Application to Two-Phase Problem

In this chapter, a two-dimensional two-phase model for stationary flow is considered on a bounded domain Ω composed of two domains Ω_1 and Ω_2. Dirichlet boundary conditions are set on $\partial\Omega$ and the estimate for the normal flux of Raviart-Thomas functions on the interpolated boundaries is applied to the interface Γ between Ω_1 and Ω_2, in order to ensure that an approximation Γ_h of this interface is sufficient to retain the optimal convergence order. Therefore, the boundary of Ω is assumed to be polygonal, as the results for isoparametric approximation combined with Dirichlet boundary are well known.

The first section is based on [26] and presents the two-phase model for a two-dimensional domain. In particular, interface conditions are introduced and a first-order formulation is derived. In the second section, the (parametric) finite-element spaces used for the approximation are presented. The corresponding least squares functional is introduced and the interface condition is estimated on the interpolated boundary. The third section uses this result in order to bound the least squares functional from below similarly to Section 5.2 and states that the optimal order of convergence is retained for the stationary two-phase problem using the (iso)parametrics element on the approximated domain. In the fourth section, a test case is discussed.

6.1 The Two-Phase Incompressible Flow Model

The domain, $\Omega \subset \mathbb{R}^2$, is assumed to be completely covered by Ω_1 and Ω_2, i.e. $\Omega = \Omega_1 \cup \Omega_2$. Further, assume that $\Omega_1 \subset \Omega_2$, let Γ denote the boundary of Ω_1 as illustrated in Figure 6.1, $\mathbf{n}$ be the unit normal on Γ that is pointing from Ω_1 to Ω_2 and κ denote the curvature of Γ. Assume that no phase transition takes place and that the two phases are viscous. Then, the two-phase model reduces to governing equations in each phase and coupling conditions at the interface. For the equations in each phase, a standard choice are the Navier-Stokes equations, as the flows are assumed to be incompressible. Further, the velocity $\mathbf{u}$ is continuous over the whole domain Ω, i.e. in the context of variational formulations $\mathbf{u}$ is searched in $H^1(\Omega)$. In opposition to the velocity, the pressure is discontinuous over the interface. Therefore $p_{\Omega_i} \in H^2(\Omega_i)$ denotes the pressure in each phase Ω_i. However, in order to simplify the notation the index i is skipped whenever the restriction on each phase is not needed. Moreover, the flows in each phase are assumed to be Newtonian, such that for the stress tensor, $\boldsymbol{\sigma}_{\Omega_i}$, it holds

$$\boldsymbol{\sigma}_{\Omega_i} = -p_{\Omega_i}\mathbf{I} + \mu_i \mathbf{D}(\mathbf{u}) \tag{6.1}$$

in each phase Ω_i, with a constant dynamic viscosity $\mu_i > 0$ and the deformation tensor $\mathbf{D}(\mathbf{u}) = \nabla\mathbf{u} + (\nabla\mathbf{u})^\top$.

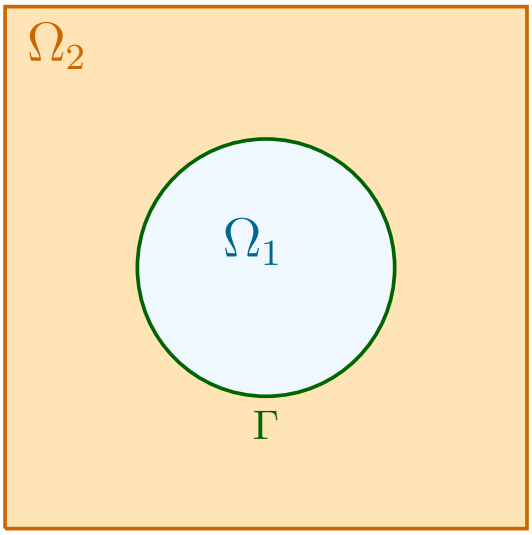


Figure 6.1: two-phase domain

Now, recall the nonstationary Navier-Stokes equations in each phase (see e.g. [26]):

$$\begin{cases} \rho_i \left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\, \mathbf{u} \right) = \operatorname{div} \boldsymbol{\sigma}_{\Omega_i} \\ \operatorname{div} \mathbf{u} = 0 \end{cases} \quad \text{in } \Omega_i,\ i = 1, 2\ , \tag{6.2}$$

where the density ρ_i in each phase is assumed to be constant. Due to the fact that on both sides of Γ there are different molecules with different attractive forces, a surface tension force acts at the interface, and this leads to the coupling condition

$$(\boldsymbol{\sigma}_{\Omega_2} - \boldsymbol{\sigma}_{\Omega_1}) \cdot \mathbf{n} = -\tau\kappa\mathbf{n} \text{ on } \Gamma\ , \tag{6.3}$$

where τ denotes the constant surface tension coefficient. Usually, time dependent problems are treated either with separated discretization in space and in time or with semi-discretization in time. In both cases, it remains to solve elliptic partial differential equations. Further, nonlinear systems are typically solved in an iterative manner, for example with Newton's method, such that the sequence of the iterates is generated recursively by solving linear systems. Therefore for the further analysis in this work, the following simplified stationary Stokes two-phase model is considered.

$$\left.\begin{aligned} \operatorname{div} \boldsymbol{\sigma}_{\Omega_i} &= \mathbf{0} \\ -p_{\Omega_i}\mathbf{I} + \mu_i \mathbf{D}(\mathbf{u}) &= \boldsymbol{\sigma}_{\Omega_i} \end{aligned}\right\} \quad \text{in } \Omega_i,\ i = 1, 2, \tag{6.4}$$

$$(\boldsymbol{\sigma}_{\Omega_2} - \boldsymbol{\sigma}_{\Omega_1}) = -\tau\kappa\mathbf{n} \qquad \text{on } \Gamma \tag{6.5}$$

Note that implies (6.4) $\operatorname{div} \mathbf{u} = 0$ and that the solution of this system is not unique, since any constant can be added on a solution to get another solution. Therefore, the normalizing condition $(p, 1)_\Omega = 0$ for the pressure has to hold. Due to (6.1) and to the symmetry of $\mathbf{D}(\mathbf{u})$, this implies $(\operatorname{tr} \boldsymbol{\sigma}, 1)_\Omega = 0$. Further, the pressure can be eliminated from the system (6.4) as well (see [33]) using the deviator of the stress tensor

$$\operatorname{dev} \boldsymbol{\sigma} = \boldsymbol{\sigma} - \frac{1}{2}(\operatorname{tr} \boldsymbol{\sigma})\mathbf{I}\ . \tag{6.6}$$

In fact, computing the deviator of the stress tensor from the second equation in (6.4) gives

$$\begin{aligned} \text{dev } \boldsymbol{\sigma}_{\Omega_i} &= -p_i\mathbf{I} + \mu_i\mathbf{D}(\mathbf{u}) - \frac{1}{2}(-2p_{\Omega_i})\mathbf{I} - \frac{1}{2}\text{tr}\left(\mu_i\mathbf{D}(\mathbf{u})\right)\mathbf{I} \\ &= \mu_i\mathbf{D}(\mathbf{u}) \ . \end{aligned} \tag{6.7}$$

This leads to the following first-order system

$$\left.\begin{aligned} \text{div } \boldsymbol{\sigma}_{\Omega_i} &= \mathbf{0} \\ \text{dev } \boldsymbol{\sigma} - \mu_i\mathbf{D}(\mathbf{u}) &= \mathbf{0} \end{aligned}\right\} \quad \text{in } \Omega,\ i=1,2, \\ (\boldsymbol{\sigma}_{\Omega_2} - \boldsymbol{\sigma}_{\Omega_1}) = -\kappa\mathbf{n} \qquad \text{on } \Gamma \\ (\text{tr } \boldsymbol{\sigma}, 1)_\Omega = 0 \tag{6.8}$$

The least squares functional associated with the problem (6.8) is

$$\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u}) = \sum_{i=1}^{2} \| \frac{1}{\sqrt{\mu_i}} \text{dev } \boldsymbol{\sigma} - \sqrt{\mu_i}\ \mathbf{D}(\mathbf{u})\|^2_{0,\Omega_i} + \|\text{div } \boldsymbol{\sigma}_{\Omega_i}\|^2_{0,\Omega_i} \tag{6.9}$$

for $\mathbf{u} \in \mathcal{W} = \left(H^1_0(\Omega)\right)^2$ and

$$\begin{aligned} \boldsymbol{\sigma} \in \boldsymbol{\Sigma} = \{\boldsymbol{\sigma} = (\boldsymbol{\sigma}_{\Omega_1}, \boldsymbol{\sigma}_{\Omega_2}) \ : \boldsymbol{\sigma}_{\Omega_i} \in (H(\text{div}, \Omega_i))^2,\ i=1,2, \\ (\boldsymbol{\sigma}_{\Omega_2} - \boldsymbol{\sigma}_{\Omega_1}) \cdot \mathbf{n} = -\kappa\mathbf{n} \text{ on } \Gamma \\ \text{and } (\text{tr } \boldsymbol{\sigma}, 1)_{0,\Omega} = 0\} \ . \end{aligned} \tag{6.10}$$

In order to simplify the notation, $[\boldsymbol{\sigma} \cdot \mathbf{n}]_\Gamma$ denotes the jump $(\boldsymbol{\sigma}_{\Omega,2} - \boldsymbol{\sigma}_{\Omega,1}) \cdot \mathbf{n}$ over Γ.

6.2 Finite-Element Spaces

Similarly to the previous work, assume that Γ is a Lipschitz continuous and piecewise C^{k+2} curve. Then, consider the interior domain Ω_1 and construct $\hat{\Omega}_1$ and a triangulation $\hat{\mathcal{T}}_{h,1}$ as in Chapter 4. Note that therefore, Γ is linearly interpolated by $\hat{\Gamma}$. Define $\hat{\Omega}_2 = \Omega \backslash \hat{\Omega}_1$ and construct a triangulation $\hat{\mathcal{T}}_{h,2}$. Note that then Ω is completely covered by both of the triangulations $\mathcal{T}_h = \mathcal{T}_{h,1} \cup \mathcal{T}_{h,2}$ and $\hat{\mathcal{T}}_h = \hat{\mathcal{T}}_{h,1} \cup \hat{\mathcal{T}}_{h,2}$. Moreover, in the higher-order case, the interpolation points of Γ for the construction of Γ_h match as their construction only depends on $\hat{\Gamma}$. Then, in both domains Ω_i, the mapping $\hat{\boldsymbol{\Phi}}_i, \boldsymbol{\Phi}_i, F_{h,i}$ can be computed conserving the same notation as in Chapter 4 with an additional index i.

As the velocity is continuous over the whole domain, standard (isoparametric) conforming elements are used for its approximation $\mathbf{u}_h$:

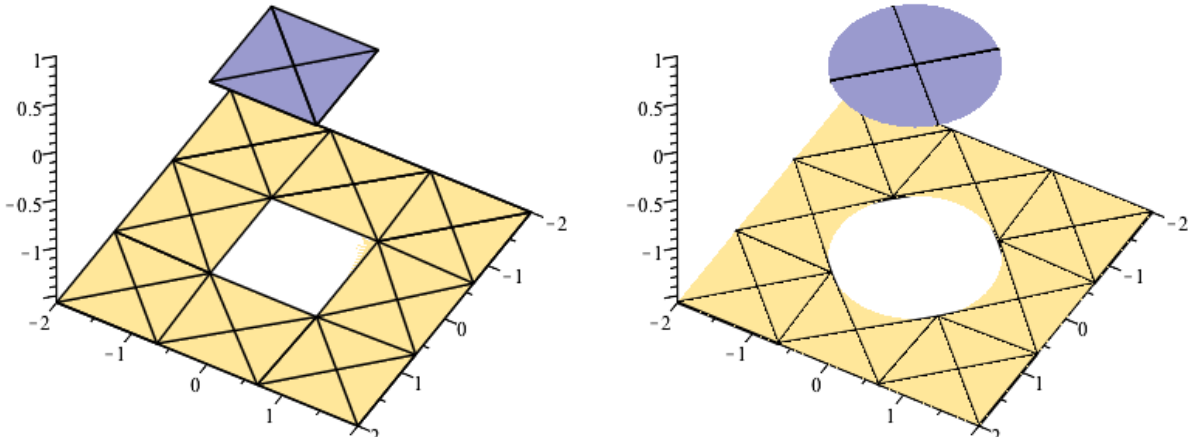
$$\mathbf{u}_h \in \mathbf{W}^{k+1}_{h,0} = \{\mathbf{w}_h = (w_{h,1}, w_{h,2}) \ : \ w_{h,i} \in Q^{k+1}_h(\Omega_i) \text{ and } w_{h,2} = 0 \text{ on } \partial\Omega_2\} \ . \tag{6.11}$$

For the approximation of the stress tensor, parametric Raviart-Thomas elements are used as each line $[\boldsymbol{\sigma}_h]_j$ of the stress tensor belongs to $H_{\text{div}}(\Omega_1) \times H_{\text{div}}(\Omega_2)$:

$$\underline{\underline{RT}}_k(\Omega) = \{(([\boldsymbol{\sigma}_{h,\Omega_1}]_1, [\boldsymbol{\sigma}_{h,\Omega_1}]_2), [\boldsymbol{\sigma}_{h,\Omega_2}]_1, [\boldsymbol{\sigma}_{h,\Omega_2}]_2)) \in \left(\widehat{RT}_k(\Omega_1)\right)^2 \times \left(\widehat{RT}_k(\Omega_2)\right)^2\} \ , \tag{6.12}$$

with the parametric Raviart-Thomas space

$$\widehat{RT}_k(\Omega_i) = \{\mathbf{v}_h : \Omega \to \mathbb{R}^2 : \mathbf{v}_h = \left(\frac{1}{\det J_{F_{h,i}}} J_{F_{h,i}} \hat{\mathbf{v}}_h\right) \circ F^{-1}_{h,i} \text{ with } \hat{\mathbf{v}}_h \in \hat{RT}_k(\Omega_i)\} \ . \tag{6.13}$$



(a) $[\grave{\mathbf{v}}_h]_1$ (b) $[\mathbf{v}_h]_1$

Figure 6.2: Location of the jump on examplary triangulation

Note that in this definition, the polynomial definition of the finite-element on $\Omega_{h,i}$ is extended in order to define those on Ω_i, similarly to Chapter 4. Hence, $\mathbf{v}_h$ jumps over Γ as illustrated in figures 6.2 and 6.3. For better distinction let $\grave{\mathbf{v}}_h$ denote the standard parametric Raviart-Thomas function that jumps on Γ_h. Further, for $[\boldsymbol{\sigma}_{h,\Omega_i}]_j \in \widehat{RT}_k(\Omega_i)$, $[\hat{\boldsymbol{\sigma}}_{h,\Omega_i}]_j \in \hat{RT}_k(\Omega_i)$ denotes the standard Raviart-Thomas function such that

$$[\boldsymbol{\sigma}_{h,\Omega_i}]_j = \left(\frac{1}{\det J_{F_{h,i}}} J_{F_{h,i}} [\hat{\boldsymbol{\sigma}}_{h,\Omega_i}]_j \right) \circ F_{h,i}^{-1} \tag{6.14}$$

Moreover, define

$$\hat{\boldsymbol{\sigma}}_h = (\hat{\boldsymbol{\sigma}}_{h,\Omega_1}, \hat{\boldsymbol{\sigma}}_{h,\Omega_2}) = (([\hat{\boldsymbol{\sigma}}_{h,\Omega_1}]_1, [\hat{\boldsymbol{\sigma}}_{h,\Omega_1}]_2), [\hat{\boldsymbol{\sigma}}_{h,\Omega_2}]_1, [\hat{\boldsymbol{\sigma}}_{h,\Omega_2}]_2)) \ . \tag{6.15}$$

In order to add the interface condition into the space $\underline{\underline{RT}}_k(\Omega)$, $\mathbf{g} = -\kappa\mathbf{n}$ has to be approximated by $\mathbf{g}_h$. Due to Theorem 1.8, for $\boldsymbol{\sigma}_h \in \underline{\underline{RT}}_k(\Omega)$ it holds

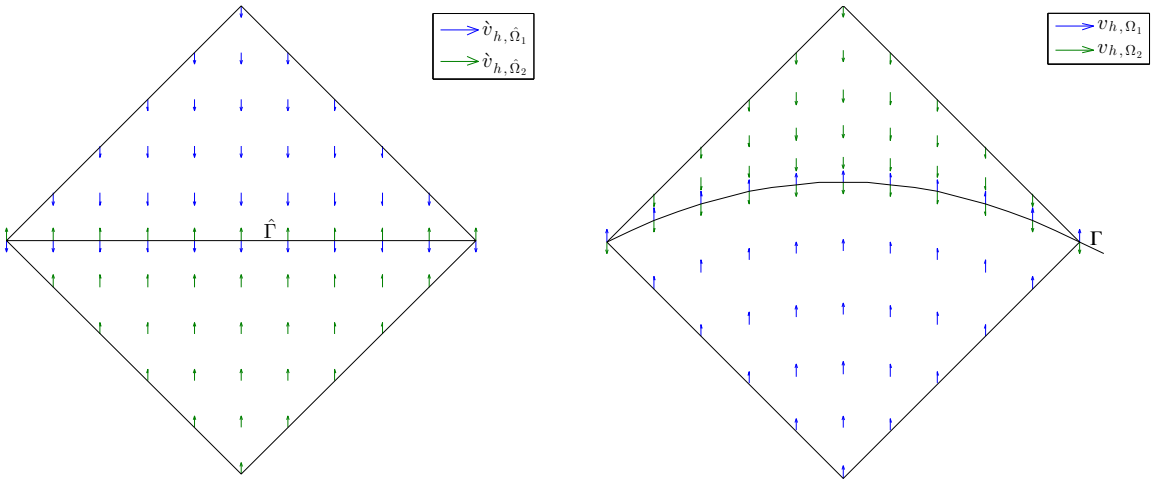
$$\grave{\boldsymbol{\sigma}}_{h,\Omega_1} \cdot \mathbf{n}|_{\Gamma_h} = \left. \frac{\hat{\boldsymbol{\sigma}}_{h,\Omega_1} \cdot \mathbf{n}}{|J_{F_{h,1}}^{-\top} \mathbf{n}| \ (\det J_{F_{h,1}})} \right|_{\Gamma_h} . \tag{6.16}$$

This motivates the definition $\mathbf{g}_h = \mathcal{Q}_h(\omega_h(\mathbf{g} \circ \hat{\boldsymbol{\Phi}}_{h,1}))$ with $\omega_h = |J_{\boldsymbol{\Psi}_{h,1}}^{\top} \mathbf{n}| \, (\det J_{F_{h,1}})$ on Γ_h, where $\mathcal{Q}_h$ is the orthogonal projection in $L^2(F_{h,1}^{-1}(\Gamma))$ onto the piecewise polynomials of degree k. Now, define

$$\boldsymbol{\Sigma}_h^k = \{\boldsymbol{\sigma}_h \in \underline{\underline{RT}}_k(\Omega) \ : \ [\grave{\boldsymbol{\sigma}}_h \cdot \mathbf{n}]_{\Gamma_h} = \mathbf{g}_h \text{ and } (\operatorname{tr} \boldsymbol{\sigma}_h, 1)_\Omega = 0\} \tag{6.17}$$

The minimizing problem corresponding to (6.8) is now given by

$$\begin{aligned} &\text{Find } (\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \boldsymbol{\Sigma}_h^k \times \mathbf{W}_{h,0}^{k+1} \text{ such that} \\ &\qquad \mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) \leq \mathcal{F}(\boldsymbol{\tau}_h, \mathbf{w}_h) \quad \forall \ (\boldsymbol{\tau}_h, \mathbf{w}_h) \in \boldsymbol{\Sigma}_h^k \times \mathbf{W}_{h,0}^{k+1} \ . \end{aligned} \tag{6.18}$$



(a) $\hat{\mathbf{v}}_h$ (b) $\mathbf{v}_h$

Figure 6.3: Example for the location of the jump for functions in $\underline{\underline{RT}}_k(\Omega)$

In the following theorem, an estimate for the jump of the normal flux on the interpolated boundary similary to Theorem 4.3 is derived. It involves the strip S_h which consists of all triangles in $\mathcal{T}_h$ whose intersection with Γ is not empty.

Theorem 6.1. *Let Ω be the two-phase domain having the properties described in Section 6.1. In particular, assume that the interface Γ is a piecewise C^{k+2} curve, $k \geq 0$. Then,*

$$|\langle [\boldsymbol{\sigma}_h \cdot \mathbf{n}]_\Gamma - \mathbf{g}, \mathbf{q} \rangle_{0,\Gamma}| \lesssim \left(h^{2k+2} \|\mathbf{g}\|_{k+1,\Gamma} + h^{2k+1} \|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)} \right)^{\frac{1}{2}} \|\mathbf{q}\|_{0,\Gamma} \tag{6.19}$$

holds for all $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h^k$ and $\mathbf{q} \in (H^{\frac{1}{2}}(\Gamma))^2$.

Proof: First, note that due to Theorem 1.8, for $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h^k$ it holds

$$[\boldsymbol{\sigma}_h \cdot \mathbf{n}]_\Gamma = \left. \frac{(J^{-\top}_{\hat{\boldsymbol{\Psi}}_{h,1}} \mathbf{n}) \cdot (J_{F_{h,1}} [\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}})}{|J^{-\top}_{\hat{\boldsymbol{\Psi}}_{h,1}} \mathbf{n}| \, (\det J_{F_{h,1}})} \right|_{\hat{\Gamma}} . \tag{6.20}$$

This implies

$$\|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]_\Gamma - \mathbf{g}\|_{0,\Gamma} = \left\| \frac{(J^{-\top}_{\hat{\boldsymbol{\Psi}}_{h,1}} \hat{\mathbf{n}}) \cdot (J_{F_{h,1}} [\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}})}{\omega_h} \circ \hat{\boldsymbol{\Psi}}_{h,1} - \mathbf{g} \right\|_{0,\Gamma} , \tag{6.21}$$

where $\hat{\mathbf{n}}$ is the normal vector on $\hat{\Gamma}$. Then, one obtains

$$\begin{aligned}\|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]_\Gamma - \mathbf{g}\|_{0,\Gamma} &\lesssim \left\| \frac{(J_{\hat{\boldsymbol{\Psi}}_{h,1}}^\top \hat{\mathbf{n}}) \cdot (J_{F_{h,1}}[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}})}{\omega_h} - \mathbf{g} \circ \hat{\boldsymbol{\Phi}}_{h,1} \right\|_{0,F_{h,1}^{-1}(\Gamma)} \\ &\leq \left\| \frac{((J_{\hat{\boldsymbol{\Psi}}_{h,1}} J_{F_{h,1}} - I)[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}})}{\omega_h} \cdot \hat{\mathbf{n}} \right\|_{0,F_{h,1}^{-1}(\Gamma)} + \left\| \frac{[\hat{\boldsymbol{\sigma}}_h \cdot \hat{\mathbf{n}}]_{\hat{\Gamma}}}{\omega_h} - \mathbf{g} \circ \hat{\boldsymbol{\Phi}}_{h,1} \right\|_{0,F_{h,1}^{-1}(\Gamma)} \\ &= \left\| \frac{((J_{\boldsymbol{\Psi}_{h,1}} - I)[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}}) \cdot \hat{\mathbf{n}}}{\omega_h} \right\|_{0,F_{h,1}^{-1}(\Gamma)} + \left\| \frac{g_h}{\omega_h} - \mathbf{g} \circ \hat{\boldsymbol{\Phi}}_{h,1} \right\|_{0,F_{h,1}^{-1}(\Gamma)}\end{aligned} \tag{6.22}$$

and thus using the fact that ω_h belongs to W_∞^{k+1},

$$\|[\boldsymbol{\sigma}_h \cdot \mathbf{n}]_\Gamma - \mathbf{g}\|_{0,\Gamma} \lesssim \left\|(J_{\boldsymbol{\Psi}_{h,1}} - I)[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}}\right\|_{0,F_{h,1}^{-1}(\Omega)} + \left\|\mathbf{g}_h - \omega_h \left(\mathbf{g} \circ \hat{\boldsymbol{\Phi}}_{h,1}\right)\right\|_{0,F_{h,1}^{-1}(\Gamma)} \ . \tag{6.23}$$

On the one hand, the term $\left\|(J_{\boldsymbol{\Psi}_{h,1}} - I)[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}}\right\|_{0,F_{h,1}^{-1}(\Omega)}$ is treated using the approximation property (4.13a) leading to

$$\begin{aligned}\left\|(J_{\boldsymbol{\Psi}_{h,1}} - I)[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}}\right\|_{0,F_{h,1}^{-1}(\Gamma)} &\lesssim h^{k+1} \left\|[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}}\right\|_{0,F_{h,1}^{-1}(\Gamma)} \\ &\lesssim h^{k+1} \|\boldsymbol{\sigma}_{h,\Omega_2} - \boldsymbol{\sigma}_{h,\Omega_1}\|_{0,\Gamma} \\ &\lesssim h^{k+1} \left(\|\boldsymbol{\sigma}_{h,\Omega_2}\|_{0,\Gamma} + \|\boldsymbol{\sigma}_{h,\Omega_1}\|_{0,\Gamma}\right).\end{aligned} \tag{6.24}$$

Using the scaling argument (2.41), this leads to

$$\begin{aligned}\left\|(J_{\boldsymbol{\Psi}_{h,1}} - I)[\hat{\boldsymbol{\sigma}}_h]_{\hat{\Gamma}}\right\|_{0,F_{h,1}^{-1}(\Gamma)} &\lesssim h^{k+\frac{1}{2}} \left(\|\boldsymbol{\sigma}_{h,\Omega_2}\|_{0,S_h(\Gamma)\cap\Omega_2} + \|\boldsymbol{\sigma}_{h,\Omega_1}\|_{0,S_h(\Gamma)\cap\Omega_1}\right) \\ &\lesssim h^{k+\frac{1}{2}} \|\boldsymbol{\sigma}_h\|_{0,S_h(\Gamma)} \ .\end{aligned} \tag{6.25}$$

On the other hand, the second term in (6.24) is estimated as

$$\left\|\mathbf{g}_h - \omega_h \left(\mathbf{g} \circ \hat{\boldsymbol{\Phi}}_{h,1}\right)\right\|_{0,F_{h,1}^{-1}(\Gamma)} \lesssim h^{k+1} \|\mathbf{g}\|_{k+1,\Gamma} \ , \tag{6.26}$$

using the minimization property of the orthogonal projection $\mathcal{Q}_h$ with respect to $L^2(F_{h,1}^{-1}(\Gamma))$ and $\omega_h \in W_\infty^{k+1}$ once more. Combining (6.24), (6.25) and (6.26) finishes this proof. □

6.3 Least Squares Functional and Ellipticity

In this section, a lower bound similary to Section 5.2 is derived. Therefore, the Korn inequality (see e.g. [12])

$$\exists\, C_K > 0 \text{ with } \|\mathbf{D}(\mathbf{w})\|_{0,\Omega} + \|\mathbf{w}\|_{0,\Omega} \geq C_K \|\mathbf{w}\|_{1,\Omega} \qquad \forall \mathbf{w} \in (H^1(\Omega))^2, \tag{6.27}$$

and the following Lemma, which is a two-phase version of Lemma 3.1 in [3] are introduced (see also [6] and [17]).

Lemma 6.1. *There exists a constant $C_D > 0$ such that*

$$\|\text{tr } \boldsymbol{\tau}\|_{0,\Omega} \leq C_D \left(\|[\boldsymbol{\tau} \cdot \mathbf{n}]\|_{-\frac{1}{2},\Gamma} + \|\text{dev } \boldsymbol{\tau}\|_{0,\Omega} + \sum_{i=1}^{2} \|\text{div } \boldsymbol{\tau}\|_{0,\Omega_i} \right) . \tag{6.28}$$

holds for all $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2) \in (H(\text{div}, \Omega_1))^d \times (H(\text{div}, \Omega_2))^d$ satisfying $(\text{tr } \boldsymbol{\tau}, 1)_{0,\Omega} = 0$.

Proof: Note that $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2) \in (H(\text{div}, \Omega_1))^2 \times (H(\text{div}, \Omega_2))^2$ implies $\text{tr } \boldsymbol{\tau} \in L^2(\Omega)$. Further, with the additional condition $(\text{tr } \boldsymbol{\tau}, 1)_{0,\Omega} = 0$, there exists $\mathbf{v} \in (H_0^1(\Omega))^2$ such that $\text{div } \mathbf{v} = \text{tr } \boldsymbol{\tau}$ and $\|\mathbf{v}\|_{1,\Omega} \lesssim \|\text{tr } \boldsymbol{\tau}\|_{0,\Omega}$ (see e.g. Lemma 11.2.3 in [14]). Then, it holds:

$$\begin{aligned}
\|\text{tr } \boldsymbol{\tau}\|_{0,\Omega}\|\mathbf{v}\|_{1,\Omega} &\lesssim \|\text{tr } \boldsymbol{\tau}\|^2 = (\text{tr } \boldsymbol{\tau}, \text{tr } \boldsymbol{\tau})_{0,\Omega} = (\text{tr } \boldsymbol{\tau}, \text{div } \mathbf{v})_{0,\Omega} \\
&= (\boldsymbol{\tau}, (\text{div } \mathbf{v})\mathbf{I})_{0,\Omega} \\
&= d\,(\boldsymbol{\tau}, \nabla\mathbf{v} - \text{dev}\,(\nabla\mathbf{v}))_{0,\Omega} \\
&= -d\,(\text{dev } \boldsymbol{\tau}, \nabla\mathbf{v})_{0,\Omega} + d\langle[\boldsymbol{\tau} \cdot \mathbf{n}], \mathbf{v}\rangle_\Gamma - d\sum_{i=1}^{2} (\text{div } \boldsymbol{\tau}, \mathbf{v})_{0,\Omega_i} \\
&\leq d\|\text{dev } \boldsymbol{\tau}\|_{0,\Omega}\|\nabla\mathbf{v}\|_{0,\Omega} + d\|\mathbf{v}\|_{\frac{1}{2},\Gamma} \sup_{\mathbf{w}\in H_0^1(\Omega)} \frac{\langle[\boldsymbol{\tau} \cdot \mathbf{n}], \mathbf{w}\rangle_\Gamma}{\|\mathbf{w}\|_{\frac{1}{2},\Gamma}} + d\sum_{i=1}^{2} \|\text{div } \boldsymbol{\tau}\|_{0,\Omega_i}\|\mathbf{v}\|_{0,\Omega_i} \\
&\leq d\|\text{dev } \boldsymbol{\tau}\|_{0,\Omega}\|\mathbf{v}\|_{1,\Omega} + d\|\mathbf{v}\|_{\frac{1}{2},\Gamma}\|[\boldsymbol{\tau} \cdot \mathbf{n}]\|_{-\frac{1}{2},\Gamma} + d\|\mathbf{v}\|_{0,\Omega}\sum_{i=1}^{2} \|\text{div } \boldsymbol{\tau}\|_{0,\Omega_i} \\
&\leq d\|\text{dev } \boldsymbol{\tau}\|_{0,\Omega}\|\mathbf{v}\|_{1,\Omega} + d\|\mathbf{v}\|_{1,\Omega}\|[\boldsymbol{\tau} \cdot \mathbf{n}]\|_{-\frac{1}{2},\Gamma} + d\|\mathbf{v}\|_{1,\Omega}\sum_{i=1}^{2} \|\text{div } \boldsymbol{\tau}\|_{0,\Omega_i} ,
\end{aligned}$$

where the trace theorem is used in the last inequality. Dividing by $\|\mathbf{v}\|_{1,\Omega}$ leads to the result. □
Using Lemma 6.1 leads now to a kind of lower bound for the least squares functional corresponding to (6.8). This is the statement of the following lemma.

Lemma 6.2. *Let $(\boldsymbol{\sigma}, \mathbf{u}) \in \boldsymbol{\Sigma} \times \mathcal{W}$ be the exact solution of the stationary two-phase problem (6.8). Then, it holds*

$$\begin{aligned}
&\mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) + \left(h^{2k+2}\|\mathbf{g}\|_{k+1,\Gamma} + h^{2k+1}\|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)}\right)^{\frac{1}{2}} \|\mathbf{u} - \mathbf{u}_h\|_{0,\Gamma} \\
&\qquad \gtrsim \|\text{dev}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2_{0,\Omega} + \|(\mathbf{u} - \mathbf{u}_h)\|^2_{1,\Omega} + \sum_{i=1}^{2} \|\text{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i}
\end{aligned} \tag{6.29}$$

for all $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \boldsymbol{\Sigma}_h^k \times \mathbf{W}_{h,0}^{k+1}$.

Proof: Due to the fact that the exact solution $(\boldsymbol{\sigma}, \mathbf{u})$ solves $\mathcal{F}(\boldsymbol{\sigma}, \mathbf{u}) = 0$, it holds

$$\mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) = \sum_{i=1}^{2} \left\| \frac{1}{\sqrt{\mu_i}}\text{dev}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u} - \mathbf{u}_h) \right\|^2_{0,\Omega_i} + \|\text{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} . \tag{6.30}$$

Further, note that due to the fact that the asymmetric part of $\mathbf{D}(\mathbf{u} - \mathbf{u}_h)$ is zero, it holds

$$\begin{aligned}
\|\text{dev}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h) - \mu_i\mathbf{D}(\mathbf{u} - \mathbf{u}_h)\|^2_{0,\Omega_i} &\geq \|\text{as}\left(\frac{1}{\sqrt{\mu_i}}(\text{dev}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u} - \mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\
&= \|\text{as}\left(\frac{1}{\sqrt{\mu_i}}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h - \frac{1}{2}\text{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u} - \mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\
&= \|\frac{1}{\sqrt{\mu_i}}\ \text{as}\,(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\,\|^2_{0,\Omega_i} .
\end{aligned} \tag{6.31}$$

Moreover, it holds tr $\mathbf{D}(\mathbf{u}-\mathbf{u}_h) = 2\mathrm{div}(\mathbf{u}-\mathbf{u}_h)$ and

$$\begin{aligned}\Big(\mathrm{dev}\ \boldsymbol{\tau}, \frac{1}{2}(\mathrm{tr}\ \boldsymbol{\tau})\mathbf{I}\Big)_{0,\Omega} &= \Big(\boldsymbol{\tau} - \frac{1}{2}(\mathrm{tr}\ \boldsymbol{\tau})\mathbf{I}, \frac{1}{2}(\mathrm{tr}\ \boldsymbol{\tau})\mathbf{I}\Big)_{0,\Omega} \\ &= \mathrm{tr}\ \boldsymbol{\tau}\Big(\boldsymbol{\tau}_{1,1} - \frac{1}{2}(\mathrm{tr}\ \boldsymbol{\tau}) + \boldsymbol{\tau}_{2,2} - \frac{1}{2}(\mathrm{tr}\ \boldsymbol{\tau})\Big)_{0,\Omega} \\ &= 0\end{aligned} \tag{6.32}$$

for all $\boldsymbol{\tau} \in (L^2(\Omega))^2 \times (L^2(\Omega))^2$. Then, using the definition of the deviatoric part on the term $\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \mu_i\mathbf{D}(\mathbf{u}-\mathbf{u}_h)$ and the fact that the trace of the deviatoric part vanishes, this implies

$$\begin{aligned}&\|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \\ &= \|\mathrm{dev}\left(\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} + \|\frac{1}{2}\mathrm{tr}\ \left(\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\ &= \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} + \|\frac{\sqrt{\mu_i}}{2}\mathrm{tr}\ \left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\ &= \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} + \mu_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}\ ,\end{aligned} \tag{6.33}$$

such that a lower bound for the least squares functional can be given by

$$\begin{aligned}\mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) &= \sum_{i=1}^{2} \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} + \|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} \\ &\gtrsim \sum_{i=1}^{2} \|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \\ &\quad + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \\ &\gtrsim \sum_{i=1}^{2} \|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\ &\quad + \mu_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} + \frac{1}{\mu_i}\|\ \mathrm{as}\left(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\right)\|^2_{0,\Omega_i}\ .\end{aligned} \tag{6.34}$$

The further steps of this proof consist in splitting the term $2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h), \mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega}$ that appears by expanding $\|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i}$ on the other terms of the least squares functional. To this end, first weight the term of the least squares functional with positive constants $C_{i,j}$, $j = 1..3$ to be defined later.

$$\begin{aligned}\mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) \gtrsim \sum_{i=1}^{2} &C_{i,1}\|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\ &+ C_{i,2}\mu_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} + \frac{C_{i,3}}{\mu_i}\|\mathrm{as}\left(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\right)\|^2_{0,\Omega_i}\ .\end{aligned} \tag{6.35}$$

Further, note that

$$\begin{aligned}2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega_i}&\\ =&2(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\mathbf{D}(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}-(\mathrm{tr}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\mathbf{I},\mathbf{D}(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}\\ =&2(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\mathbf{D}(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}-2(\mathrm{tr}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{div}(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}\\ =&2(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h,\nabla(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}-2(\mathrm{as}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{as}(\nabla(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega_i}\\ &-2(\mathrm{tr}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{div}(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}\end{aligned}$$

holds, where $\mathbf{D}(\mathbf{u}-\mathbf{u}_h)$ was splitted to its symmetric and asymmetric parts in the last inequality. Integrating by parts one obtains

$$\begin{aligned}2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),&\mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega}\\ =&2\langle[(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\cdot\mathbf{n}],(\mathbf{u}-\mathbf{u}_h)\rangle_{0,\Gamma}+\sum_{i=1}^{2}-2(\mathrm{as}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{as}\,\nabla(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}\\ &-(\mathrm{tr}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{div}(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}-2(\mathrm{div}((\boldsymbol{\sigma}_{\Omega_i}-\boldsymbol{\sigma}_{h,\Omega_i}),(\mathbf{u}-\mathbf{u}_h))_{0,\Omega_i}\end{aligned}$$

such that

$$\begin{aligned}-|2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),&\mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega}|\\ \geq&-|2\langle\mathbf{g}-[\boldsymbol{\sigma}_h\cdot\mathbf{n}],(\mathbf{u}-\mathbf{u}_h)\rangle_{0,\Gamma}|-\sum_{i=1}^{2}\frac{1}{\alpha_i}\|\mathrm{as}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i}+\alpha_i\|\mathrm{as}\,\nabla(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}\\ &\frac{1}{\beta_i}\|\mathrm{tr}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i}+\beta_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}+\alpha_i\|\mathbf{u}-\mathbf{u}_h\|^2_{0,\Omega_i}\\ &+\frac{1}{\alpha_i}\|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i}-\boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i}\ ,\end{aligned}$$

holds for any $\alpha_i,\beta_i>0$. Using the inequality (6.19) and the Korn inequality (6.27) leads to

$$\begin{aligned}-|2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),&\mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega}|\\ \gtrsim&-\left(h^{2k+2}\|\mathbf{g}\|_{k+1,\Gamma}+h^{2k+1}\|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)}\right)^{\frac{1}{2}}\|\mathbf{u}-\mathbf{u}_h\|_{0,\Gamma}\\ &-\sum_{i=1}^{2}\frac{1}{\alpha_i}\|\mathrm{as}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i}+\beta_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}+2\frac{\alpha_i}{C_K}\|\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}\\ &+\frac{C_D}{\beta_i}\|\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i}+\left(\frac{C_D}{\beta_i}+\frac{1}{\alpha_i}\right)\|\mathrm{div}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i}\end{aligned}\tag{6.36}$$

where the Lemma 6.1 was used to obtain the last inequality. Combining this with the fact that

$$\begin{aligned}\|\mathrm{dev}\,(\mathbf{D}(\mathbf{u}-\mathbf{u}_h))\,\|^2_{0,\Omega_i}=&\|\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}+\|\frac{1}{2}\mathrm{tr}\ (\mathbf{D}(\mathbf{u}-\mathbf{u}_h))\,\mathbf{I}\|^2_{0,\Omega_i}\\ &-(\mathbf{D}(\mathbf{u}-\mathbf{u}_h),\mathrm{tr}\ (\mathbf{D}(\mathbf{u}-\mathbf{u}_h))\mathbf{I})_{0,\Omega_i}\\ =&\|\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}+2\|\mathrm{div}\,(\mathbf{u}-\mathbf{u}_h)\,\|^2_{0,\Omega_i}\\ &-2\,(\mathbf{D}(\mathbf{u}-\mathbf{u}_h),\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\mathbf{I})_{0,\Omega_i}\\ =&\|\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i}-2\|\mathrm{div}\,(\mathbf{u}-\mathbf{u}_h)\,\|^2_{0,\Omega_i}\end{aligned}\tag{6.37}$$

holds and expanding the mixed term in (6.35) leads to

$$\begin{aligned}
\mathcal{F}(\boldsymbol{\sigma}_h,\mathbf{u}_h) \gtrsim & \sum_{i=1}^{2} C_{i,1}\|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i}-\boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h) - \sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} \\
& + C_{i,2}\mu_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} + \frac{C_{i,3}}{\mu_i}\|\mathrm{as}\left(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\right)\|^2_{0,\Omega_i} \\
\gtrsim & \sum_{i=1}^{2} C_{i,1}\|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i}-\boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i} \\
& + \|\sqrt{\mu_i}\mathrm{dev}\left(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\right)\|^2_{0,\Omega_i} \quad + C_{i,2}\mu_i\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \\
& + \frac{C_{i,3}}{\mu_i}\|\mathrm{as}\left(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\right)\|^2_{0,\Omega_i} - 2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega} \\
\gtrsim & \sum_{i=1}^{2} C_{i,1}\|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i}-\boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + \frac{1}{\mu_i}\|\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i} + \mu_i\|\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \\
& + \mu_i C_{i,2}\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} + \frac{C_{i,3}}{\mu_i}\|\mathrm{as}\left(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\right)\|^2_{0,\Omega_i} \\
& - 2(\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h),\mathrm{dev}(\mathbf{D}(\mathbf{u}-\mathbf{u}_h)))_{0,\Omega} \\
\gtrsim & \sum_{i=1}^{2}(C_{i,1}-\frac{C_D}{\beta_i}-\frac{1}{\alpha_i})\|\mathrm{div}(\boldsymbol{\sigma}_{\Omega_i}-\boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i} + (\frac{1}{\mu_i}-\frac{C_D}{\beta_i})\|\mathrm{dev}(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h)\|^2_{0,\Omega_i} \\
& + (\mu_i - 2\frac{\alpha_i}{C_K})\|\mathbf{D}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \quad + (\mu_i C_{i,2}-\beta_i)\|\mathrm{div}(\mathbf{u}-\mathbf{u}_h)\|^2_{0,\Omega_i} \\
& + (\frac{C_{i,3}}{\mu_i}-\frac{1}{\alpha_i})\|\mathrm{as}\left(\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\right)\|^2_{0,\Omega_i} \\
& - \left(h^{2k+2}\|\mathbf{g}\|_{k+1,\Gamma} + h^{2k+1}\|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)}\right)^{\frac{1}{2}}\|\mathbf{u}-\mathbf{u}_h\|_{0,\Gamma}
\end{aligned} \tag{6.38}$$

Choosing

$$\alpha_i = \frac{1}{4}C_K\mu_i, \qquad C_{i,3} = \frac{\mu_i}{\alpha_i}, \qquad \beta_i = 2\mu_i C_D,$$
$$C_{i,2} = 3C_D, \qquad C_{i,1} = \frac{1}{\mu_i}$$

and using the Korn inequality once more finishes the proof. □

The following theorem states that the optimal order of convergence is retained for the stationary two-phase problem (6.8) using the (iso)parametric element on the approximated domain Ω_h, i.e. by using the approximated least squares functional

$$\mathcal{F}_h(\boldsymbol{\sigma},\mathbf{u}) = \sum_{i=1}^{2}\|\mathrm{div}\ \boldsymbol{\sigma}_{\Omega_i}\|^2_{0,\Omega_{h,i}} + \|\frac{1}{\sqrt{\mu_i}}\mathrm{dev}\ \boldsymbol{\sigma} - \sqrt{\mu_i}\mathbf{D}(\mathbf{u})\|^2_{0,\Omega_{h,i}}\ . \tag{6.39}$$

Theorem 6.2. *Let $(\boldsymbol{\sigma},\mathbf{u}) \in \boldsymbol{\Sigma}_h^k \times \mathcal{W}$ denote the exact solution of the system (6.8) and assume that it satisfies $\mathbf{u} \in (H^{k+2}(\Omega))^2$ and $\mathrm{div}\ \boldsymbol{\sigma}_i \in (H^{k+1}(\Omega_i))^4$. Further, let $(\boldsymbol{\sigma}_h,\mathbf{u}_h) \in \boldsymbol{\Sigma}_h^k \times \mathbf{W}_{h,0}^{k+1}$ denote the (parametric) finite-element approximation minimizing $\mathcal{F}_h(\boldsymbol{\sigma}_h,\ \mathbf{u}_h)$ under all $(\boldsymbol{\tau}_h,\mathbf{w}_h) \in \boldsymbol{\Sigma}_h^k \times \mathbf{W}_{h,0}^{k+1}$. Then it holds*

$$\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_h\|_{\mathrm{div},\Omega} + \|\mathbf{u}-\mathbf{u}_h\|_{1,\Omega} \lesssim h^{k+1}\left(\|\mathbf{u}\|_{k+2,\Omega} + \sum_{i=1}^{2}\|\boldsymbol{\sigma}_i\|_{k+1,\Omega_i} + |\mathrm{div}\ \boldsymbol{\sigma}_i|_{k+1,\Omega_i}\right)\ . \tag{6.40}$$

Proof: The proof is similar to the proof of Theorem 5.1. However, the proof of the equivalence between $\mathcal{F}$ and $\mathcal{F}_h$ on the finite-dimensional spaces simplifies in this case since the two-phase problem has no right-hand side. Equation (5.27) reduces to

$$\begin{aligned}\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h) &\lesssim \sum_{i=1}^{2} \|\text{div } \boldsymbol{\sigma}_{h,\Omega_i}\|^2_{0,\Omega_{h,i}} + \|\frac{1}{\sqrt{\mu_i}}\text{dev } \boldsymbol{\sigma}_h - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}_h)\|^2_{0,\Omega_{h,i}} \\ &\lesssim \sum_{i=1}^{2} \|\text{div } \boldsymbol{\sigma}_{h,\Omega_i}\|^2_{0,\Omega_i} + \|\frac{1}{\sqrt{\mu_i}}\text{dev } \boldsymbol{\sigma}_h - \sqrt{\mu_i}\mathbf{D}(\mathbf{u}_h)\|^2_{0,\Omega} \\ &\lesssim \mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) \ ,\end{aligned} \tag{6.41}$$

using the fact that the norms on finite dimensional spaces are equivalent. The second step is similar to the proof of Lemma 5.3. The exact solution is mapped on Ω_h using the mapping $\boldsymbol{\Phi}_h : \Omega_h \to \Omega$ defined in (4.11):

$$\widetilde{\boldsymbol{\sigma}} = \left[\left(\frac{1}{\det J_{\boldsymbol{\Psi}_h}} J_{\boldsymbol{\Psi}_h}[\boldsymbol{\sigma}]_1\right) \left(\frac{1}{\det J_{\boldsymbol{\Psi}_h}} J_{\boldsymbol{\Psi}_h}[\boldsymbol{\sigma}]_2\right)\right]^\top \circ \boldsymbol{\Phi}_h \tag{6.42}$$

$$\widetilde{\mathbf{u}} = \mathbf{u} \circ \boldsymbol{\Phi}_h \ . \tag{6.43}$$

Due to Theorem 1.8, it holds

$$\text{div } \widetilde{\boldsymbol{\sigma}} = \left[\frac{1}{\det J_{\boldsymbol{\Psi}_h}} \text{div } [\boldsymbol{\sigma}]_1 \ \frac{1}{\det J_{\boldsymbol{\Psi}_h}} \text{div } [\boldsymbol{\sigma}]_2\right]^\top \circ \boldsymbol{\Phi}_h \tag{6.44}$$

$$\nabla\widetilde{\mathbf{u}} = \left(J^T_{\boldsymbol{\Phi}_h} \nabla \mathbf{u}\right) \circ \boldsymbol{\Phi}_h \ . \tag{6.45}$$

Now, define the interpolation operators from Section 4.2 for each components, i.e.

$$\mathcal{R}_h\boldsymbol{\sigma} = [\mathcal{R}_h([\boldsymbol{\sigma}]_1) \ \mathcal{R}_h([\boldsymbol{\sigma}]_2)]^\top \tag{6.46}$$

and

$$\mathcal{I}_h\mathbf{u} = [\mathcal{I}_h([\mathbf{u}]_1) \ \mathcal{I}_h([\mathbf{u}]_2)]^\top \ . \tag{6.47}$$

Using the fact that $(\boldsymbol{\sigma}_h, \mathbf{u}_h) \in \boldsymbol{\Sigma}^k_h \times \mathbf{W}^{k+1}_{h,0}$ minimizes $\mathcal{F}_h(\boldsymbol{\tau}_h, \ \mathbf{w}_h)$ under all $(\boldsymbol{\tau}_h, \mathbf{w}_h) \in \boldsymbol{\Sigma}^k_h \times \mathbf{W}^{k+1}_{h,0}$, leads to

$$\begin{aligned}F(\boldsymbol{\sigma}_h, \mathbf{u}_h) &\leq F_h(\mathcal{R}_h\widetilde{\boldsymbol{\sigma}}, \mathcal{I}_h\widetilde{\mathbf{u}}) \\ &= \|\frac{1}{\sqrt{\mu_i}}\text{dev } \mathcal{R}_h\widetilde{\boldsymbol{\sigma}} - \sqrt{\mu_i}\mathbf{D}(\mathcal{I}_h\widetilde{\mathbf{u}})\|^2_{0,\Omega} + \sum_{i=1}^{2} \|\text{div } \mathcal{R}_h\widetilde{\boldsymbol{\sigma}}_{\Omega_i}\|^2_{0,\Omega_{h,i}} \\ &= \sum_{i=1}^{2} \|\text{div } \mathcal{R}_h\widetilde{\boldsymbol{\sigma}}_{\Omega_i} - \text{div } \widetilde{\boldsymbol{\sigma}}_{\Omega_i} + \text{div } \widetilde{\boldsymbol{\sigma}}_{\Omega_i}\|^2_{0,\Omega_{h,i}} \\ &\quad + \|\frac{1}{\sqrt{\mu_i}}\left(\text{dev } \mathcal{R}_h\widetilde{\boldsymbol{\sigma}} - \text{dev } \widetilde{\boldsymbol{\sigma}} + \text{dev } \widetilde{\boldsymbol{\sigma}}\right) - \sqrt{\mu_i}\mathbf{D}(\widetilde{\mathbf{u}}) + \sqrt{\mu_i}\left(\mathbf{D}(\widetilde{\mathbf{u}}) - \mathbf{D}(\mathcal{I}_h\widetilde{\mathbf{u}})\right)\|^2_{0,\Omega} \\ &\lesssim \sum_{i=1}^{2} \|\text{div } \mathcal{R}_h\widetilde{\boldsymbol{\sigma}}_{\Omega_i} - \text{div } \widetilde{\boldsymbol{\sigma}}_{\Omega_i}\|^2_{0,\Omega_{h,i}} + \|\text{div } \widetilde{\boldsymbol{\sigma}}_{\Omega_i}\|^2_{0,\Omega_{h,i}} \\ &\quad + \|\text{dev } \mathcal{R}_h\widetilde{\boldsymbol{\sigma}} - \text{dev } \widetilde{\boldsymbol{\sigma}}\|^2_{0,\Omega} + \|\frac{1}{\sqrt{\mu_i}}\text{dev } \widetilde{\boldsymbol{\sigma}} - \sqrt{\mu_i}\mathbf{D}(\widetilde{\mathbf{u}})\|^2_{0,\Omega} + \|\mathbf{D}(\widetilde{\mathbf{u}}) - \mathbf{D}(\mathcal{I}_h\widetilde{\mathbf{u}})\|^2_{0,\Omega} \ .\end{aligned} \tag{6.48}$$

For the second term, using (4.13b) leads to

$$\|\operatorname{div} \widetilde{\boldsymbol{\sigma}}_{\Omega_i}\|^2_{0,\Omega_{h,i}} = \|\frac{1}{\det J_{\boldsymbol{\Psi}_{h,i}}}\operatorname{div} \boldsymbol{\sigma}_{\Omega_i}\|^2_{0,\Omega} \lesssim h^{2k+2}\|\boldsymbol{\sigma}_{\Omega_i}\|_{0,\Omega} \tag{6.49}$$

and similarly, for the fourth term it holds

$$\|\frac{1}{\sqrt{\mu_i}}\operatorname{dev} \widetilde{\boldsymbol{\sigma}} - \sqrt{\mu_i}\mathbf{D}(\widetilde{\mathbf{u}})\|^2_{0,\Omega} \lesssim h^{2k+2}\left(\|\operatorname{dev} \boldsymbol{\sigma}\|^2_{0,\Omega} + \|\mathbf{D}(\mathbf{u})\|^2_{0,\Omega}\right) . \tag{6.50}$$

Estimating the first, third and fifth of the terms in (6.48) by the interpolation estimates from (4.25) and (4.27) and then inserting (6.49) and (6.50) leads to

$$\mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) \lesssim h^{2k+2}\left(\|\mathbf{u}\|^2_{2,\Omega} + \sum_{i=1}^{2}\|\boldsymbol{\sigma}_{\Omega_i}\|^2_{k+1,\Omega_i} + |\operatorname{div} \boldsymbol{\sigma}_{\Omega_i}|^2_{k+1,\Omega_i}\right) \tag{6.51}$$

which combined with (6.41) implies

$$\mathcal{F}_h(\boldsymbol{\sigma}_h, \mathbf{u}_h) \lesssim h^{2k+2}\left(\|\mathbf{u}\|^2_{2,\Omega} + \sum_{i=1}^{2}\|\boldsymbol{\sigma}_{\Omega_i}\|^2_{k+1,\Omega_i} + |\operatorname{div} \boldsymbol{\sigma}_{\Omega_i}|^2_{k+1,\Omega_i}\right) . \tag{6.52}$$

Further, combining this with Lemmas 6.1 and 6.2, this leads to

$$\begin{aligned}
\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|^2_{0,\Omega} &\lesssim \|\operatorname{dev}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2_{0,\Omega} + \|\operatorname{tr}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2_{0,\Omega} \\
&\lesssim \left(h^{2k+2}\|\mathbf{g}\|_{k+1,\Gamma} + h^{2k+1}\|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)}\right) \\
&\qquad + h^{2k+2}\left(\|\mathbf{u}\|^2_{2,\Omega} + \sum_{i=1}^{2}\|\boldsymbol{\sigma}_{\Omega_i}\|^2_{k+1,\Omega_i} + |\operatorname{div} \boldsymbol{\sigma}_{\Omega_i}|^2_{k+1,\Omega_i}\right) \\
&\lesssim h\left(\|\mathbf{g}\|_{k+1,\Gamma} + \|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)}\right) .
\end{aligned} \tag{6.53}$$

The next step consists in applying the following estimate (see [28], Lemma 2.1)

$$\|\boldsymbol{\sigma}\|_{0,S_h(\Gamma)\cap\Omega_i} \leq h^{\frac{1}{2}}\|\boldsymbol{\sigma}\|_{1,\Omega_i} \tag{6.54}$$

to retain the optimal convergence order:

$$\frac{1}{h}\|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)\cap\Omega_i} \leq \frac{1}{h}\|\boldsymbol{\sigma}\|^2_{0,S_h(\Gamma)\cap\Omega_i} + \frac{1}{h}\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|^2_{0,S_h(\Gamma)\cap\Omega_i} \tag{6.55}$$

$$\lesssim \|\boldsymbol{\sigma}\|^2_{1,\Omega_i} + \frac{1}{h}\|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|^2_{0,S_h(\Gamma)\cap\Omega_i} . \tag{6.56}$$

Combine this with (6.53) leads to

$$\frac{1}{h}\|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)\cap\Omega_i} \lesssim \|\mathbf{g}\|_{k+1,\Gamma} + \|\boldsymbol{\sigma}_h\|^2_{0,S_h(\Gamma)}. \tag{6.57}$$

Combine this again with (6.2) leads to

$$\begin{aligned}
&\mathcal{F}(\boldsymbol{\sigma}_h, \mathbf{u}_h) + h^{k+1}\left(\|\mathbf{g}\|_{k+1,\Gamma} + \sum_{i=1}^{2}\|\boldsymbol{\sigma}\|^2_{1,\Omega_i}\right)^{\frac{1}{2}}\|\mathbf{u} - \mathbf{u}_h\|_{0,\Gamma} \\
&\qquad \gtrsim \|\operatorname{dev}(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)\|^2_{0,\Omega} + \|(\mathbf{u} - \mathbf{u}_h)\|^2_{1,\Omega} + \sum_{i=1}^{2}\|\operatorname{div}(\boldsymbol{\sigma}_{\Omega_i} - \boldsymbol{\sigma}_{h,\Omega_i})\|^2_{0,\Omega_i}
\end{aligned} \tag{6.58}$$

Finally, combining (6.52) with (6.58) in the same way as in Theorem 5.1 leads to the result. □

6.4 A Theoretical Example

A simple example of the problem (6.8) is given for a rotational flow $\mathbf{u}(\mathbf{x}) = (-y, x)$ in a two-phase domain with a unit circular interface (see e.g. [23]). Then, $\mathbf{D}(\mathbf{u}) = 0$ and thus, the deviator of the stress tensor is zero. Due to the fact that in each phase, the stress tensor is divergence-free as well, this implies $\boldsymbol{\sigma}_{\Omega_i} = \alpha_i \mathbf{I}$ for constants α_i. Then, the interface condition leads to $\alpha_2 = -\kappa + \alpha_1 = -1 + \alpha_1$. Considering the normalizing condition $(\operatorname{tr} \boldsymbol{\sigma}, 1)_\Omega = 0$ leads to

$$\alpha_1 |\Omega_1| + \alpha_2 |\Omega_2| = 0 \tag{6.59}$$

and thus to the following exact solution, that only depends on the surface of the domains Ω_i.

$$\boldsymbol{\sigma}_{\Omega_1} = \frac{|\Omega_2|}{|\Omega|} \mathbf{I} \tag{6.60}$$

$$\boldsymbol{\sigma}_{\Omega_2} = -\frac{|\Omega_1|}{|\Omega|} \mathbf{I} \,. \tag{6.61}$$

The corresponding pressure $p_i = \operatorname{tr} \boldsymbol{\sigma}_i = 2\alpha_i$ is shown for $\Omega_2 = [-2, 2] \times [-2, 2]$ in Figure 6.4.

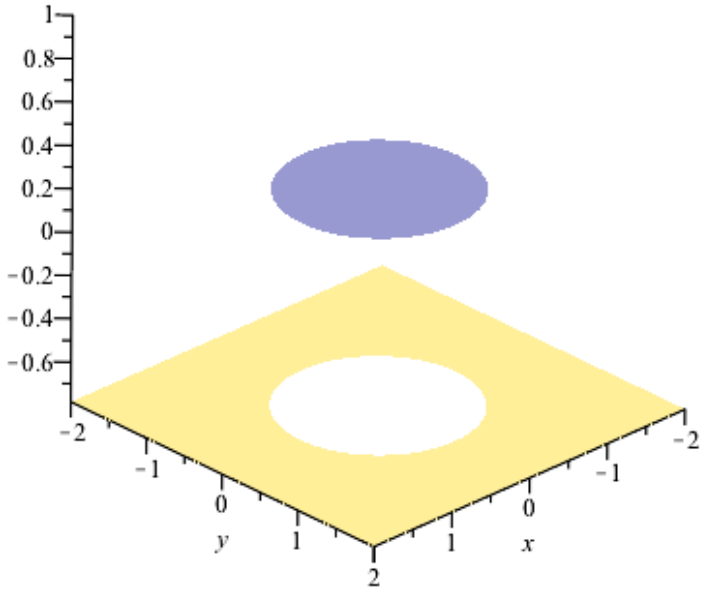


Figure 6.4: Exact pressure corresponding for the rotational flow $\mathbf{u}(\mathbf{x}) = (-y, x)$

.

This solution belong to the finite-element space considered so far. Hence, only the approximation of the domain leads to an approximation of the solution. The numerical computation

leads to the solution

$$\boldsymbol{\sigma}_{\Omega_{h,1}} = -\frac{|\Omega_{h,2}|}{|\Omega|} \tag{6.62}$$

$$\boldsymbol{\sigma}_{\Omega_{h,2}} = \frac{|\Omega_{h,1}|}{|\Omega|} \ . \tag{6.63}$$

which is as good as the approximation of the domain. In particular, this clearly demonstrates that a polygonal approximation of the boundary is not sufficient to retain the optimal convergence order in the higher-order case.

Chapter 7

Considerations for Three-Dimensional Problems

In this chapter, a bounded domain $\Omega \subset \mathbb{R}^3$ with a piecewise C^{k+2} and Lipschitz continuous boundary Γ is considered. The construction of the corresponding polygonal domain $\hat{\Omega}$ described in Chapter 2 cannot be extended to the three-dimensional case. Therefore, the aim of the first section in this chapter is to present the procedure from [27] for the construction of the domain $\hat{\Omega}$ and of the mapping $\mathbf{\Phi} : \hat{\Omega} \to \Omega$, which satisfies the conditions of Lemma 2.1. The second section introduces the Lagrange finite-elements for three-dimensional domains and the third section considers the Raviart-Thomas finite-elements. In the fourth section numerical results on the three-dimensional unit ball are presented.

7.1 Construction of the Approximated Domain

The first step of the construction of the approximated domain is the interpolation of the boundary. Therefore, let $\boldsymbol{\gamma} \ : \ D \subset \mathbb{R}^2 \to \mathbb{R}^3$ denote a parametrization of Γ such that

$$\Gamma = \{\boldsymbol{\gamma}(\mathbf{x}) \ : \ \mathbf{x} \in D\} \ . \tag{7.1}$$

Note that $D \subset \mathbb{R}^2$ implies that it can be triangulated by a triangulation $\mathcal{S}_h$ such that

(i) $\{\mathcal{S}_{h,p_i}\}_{i=1}^{N_P}$ denotes the set of vertices in $\mathcal{S}_h$.

(ii) The polygon $\hat{\Omega}$ formed with all the vertices $\{\boldsymbol{\gamma}_i = \boldsymbol{\gamma}(\mathcal{S}_{h,p_i})\}_{i=1}^{N_P}$ is simple. $\hat{\Gamma} = \partial\hat{\Omega}$ can be parametrized with $\hat{\boldsymbol{\gamma}} = \mathcal{I}_h\boldsymbol{\gamma}$ where $\mathcal{I}_h$ denotes the linear Lagrange interpolation operator defined in Section 1.5.1 with respect to the triangulation $\mathcal{S}_h$.

(iii) $\{\boldsymbol{\gamma}_i\}_{i=1}^{N_P}$ contains all points where Γ is not C^{k+2}.

(iv) $\mathcal{S}_h$ consists of N triangles $\mathcal{T}_i$ with vertices $\{\mathcal{T}_{i,j}\}_{j=1}^{3}$, $i = 1, ..., N$.

Now, let $\hat{\mathcal{T}}_h$ denote a quasi-uniform triangulation (consisting of $\bar{N}$ tetrahedra) of $\hat{\Omega}$ such that the boundary points of $\hat{\mathcal{T}}_h$ are given with $\{\boldsymbol{\gamma}_i\}_{i=1}^{N_P}$. Similarly to Chapter 2, for $i \leq N$, let $\hat{T}_i$ denote the tetrahedron of $\hat{\mathcal{T}}_h$ whose intersection with $\hat{\Gamma}$ is the triangle $\hat{\boldsymbol{\gamma}}(\mathcal{T}_i)$. Contrary to the two-dimensional case, for $i > N$, the intersection of the tetrahedron $\hat{T}_i$ with $\hat{\Gamma}$ is not necessary a unique point, as it can be a whole edge of $\hat{\boldsymbol{\gamma}}(\mathcal{T}_i)$, see Figure 7.1. Assume that the tetrahedra

are numbered such that dim $(\hat{T}_i \cap \hat{\Gamma}) = 1$ if and only if $N < i \leq \mathring{N}$. The further steps imply the reference tetrahedron

$$\hat{T}_{ref} = \left\{ \mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \ : \ x_i \geq 0 \text{ and } \sum_{i=0}^{3} x_i \leq 1 \right\} \tag{7.2}$$

and the reference transformation $\mathbf{F}_{\text{ref},\hat{T}_i} : \hat{T}_i \to T_{\text{ref}}$ such that

$$\mathbf{F}_{\text{ref},\hat{T}_i}(\hat{T}_i \cap \hat{\Gamma}) = \begin{cases} F_{\text{ref}} = \{(x_1, x_2, 0) \ : \ 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 - x_1\} & i \leq N \\ E_{\text{ref}} = \{(x_1, 0, 0) \ : \ 0 \leq x_1 \leq 1\} & N < i \leq \mathring{N} \end{cases} \tag{7.3}$$

holds. Moreover, the following notation is introduced:

- For $i \leq N$, $\mathscr{V}_3(\hat{T}_i)$ denotes the vertex of $\hat{T}_i$ whose intersection with $\hat{\boldsymbol{\gamma}}(\mathscr{F}_i)$ is empty.
- $\mathscr{V}_3(\hat{T}_{\text{ref}}) = (0, 0, 1)$.
- For $i \leq N$, $F_i = \hat{\boldsymbol{\gamma}}(\mathscr{F})$ is the triangle with the vertices $\{\boldsymbol{\gamma}(\mathscr{F}_{i,j})\}_{j=1}^{3}$ and $\widetilde{F}_i = \boldsymbol{\gamma}(\mathscr{F}_i)$ denotes the corresponding curved triangle on Γ_i.
- For $N < i \leq \mathring{N}$, E_i denotes the edge $\hat{T}_i \cap \hat{\Gamma}_i$ and $\widetilde{E}_i = \boldsymbol{\gamma}(E_i)$.
- For $N < i \leq \mathring{N}$, $\bar{E}_i$ denotes the edge of $\hat{T}_i$ whose intersection with E_i is empty.
- $\bar{E}_{\text{ref}} = \{(0, t, 1 - t) \ : \ t \in [0, 1]\}$.

Figure 7.1 summarizes these notations. Replacing the boundary triangle F_i of $\hat{T}_i$ by the curved triangle $\widetilde{F}_i$ leads to a curved tetrahedron $\widetilde{T}_i$ for $i \leq N$. For $N < i \leq \mathring{N}$, the curved tetrahedron $\widetilde{T}_i$ is obtained by replacing the edge E_i by the curved one $\widetilde{E}_i$.

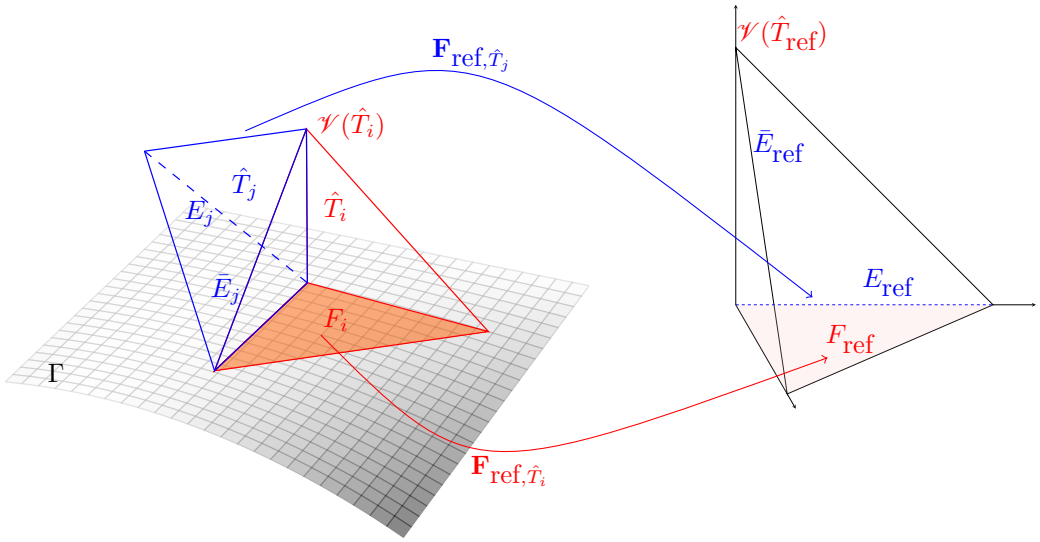


Figure 7.1: The intersection of $\hat{\Gamma}$ with $\hat{T}$ is a face (case $\hat{T}_i$) or an edge (case $\hat{T}_j$)

For $i \leq N$, the construction of the mapping $\hat{\boldsymbol{\Phi}}_h : \hat{T}_i \to \widetilde{T}_i$ is similar to the two-dimensional case. It involves a mapping $\mathbf{Z}^2$ that connects an interior point with a corresponding point on

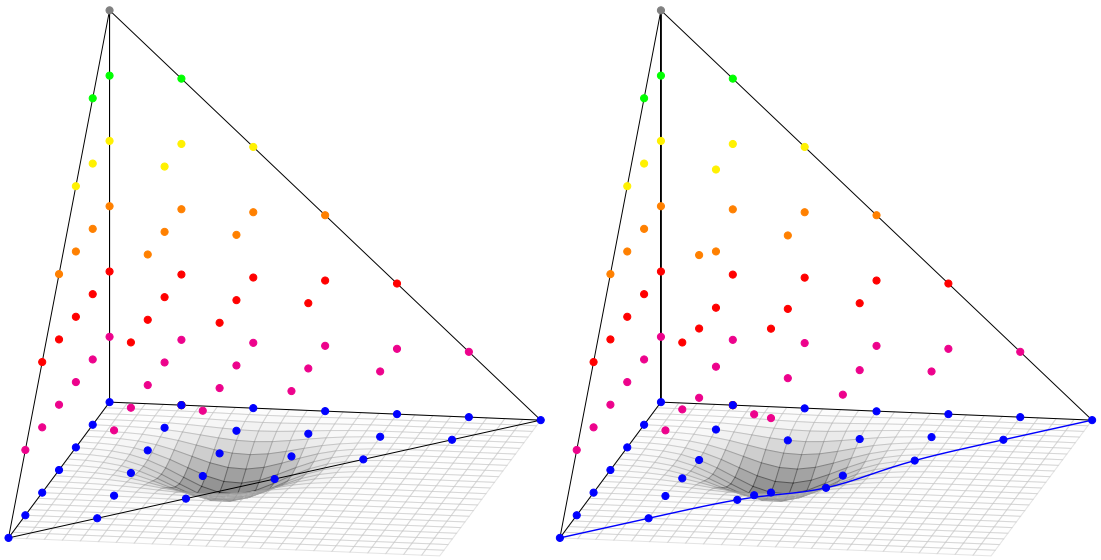


Figure 7.2: Example of the streching resulting from $\mathbf{Z}_i^2,\ i \leq N$

the approximated boundary, whose construction is similar to the construction of $\mathbf{Z}$ in Chapter 2, i.e. the mapping $\mathbf{Z}_{\text{ref}}^2$ to connect an interior point with a corresponding point on the edge F_{ref} is needed first. For a point $\mathbf{x}_{\text{ref}} = (x_{\text{ref}}, y_{\text{ref}}, z_{\text{ref}}) \in \hat{T}_{\text{ref}}$ consider the line from $(0,0,1)$ through the point $\mathbf{x}_{\text{ref}}$. As it crosses the face F_{ref} at the point $\left(-\frac{x_{\text{ref}}}{z_{\text{ref}}-1}, -\frac{y_{\text{ref}}}{z_{\text{ref}}-1}, 0\right)$ the mapping $\mathbf{Z}_{\text{ref}}$ is defined as follows:

$$\begin{aligned} \mathbf{Z}_{\text{ref}}^2 \ : \ \hat{T}_{\text{ref}} &\to \ F_{\text{ref}} \\ \mathbf{x}_{\text{ref}} &\mapsto \ \left(\frac{x_{\text{ref}}}{1-z_{\text{ref}}}, \frac{y_{\text{ref}}}{1-z_{\text{ref}}}, 0\right) \ . \end{aligned}$$

With this mapping $\mathbf{Z}_{\text{ref}}^2$, a mapping $\mathbf{Z}_i^2 = \mathbf{F}_{\text{ref},\hat{T}_i} \circ \mathbf{Z}_{\text{ref}}^2 \circ \mathbf{F}_{\text{ref},\hat{T}_i}^{-1}$ can be defined to connect an interior point of any tetrahedron $\hat{T}_i$ with a corresponding point on $\hat{\Gamma}_i$. Similarly to Chapter 2, any point $\mathbf{x} \in \hat{T}_i$ is located on the line segment $[\mathscr{V}_3(\hat{T}_i), \mathbf{Z}_i^2(\mathbf{x})]$ and the position of $\mathbf{x}$ on this line can be given with the ratio $\delta(\mathbf{x})$ of the distance between $\mathscr{V}_3(\hat{T}_i)$ and $\mathbf{x}$ to the distance between $\mathscr{V}_3(\hat{T}_i)$ and $\mathbf{Z}_i^2(\mathbf{x})$, i.e :

$$\delta(\mathbf{x}) = \frac{\text{dist}(\mathscr{V}_3(\hat{T}_i), \mathbf{x})}{\text{dist}(\mathscr{V}_3(\hat{T}_i), \mathbf{Z}_i^2(\mathbf{x}))} \tag{7.4}$$

Note that due to the fact that the affine mapping conserves the ratio of the distances, it holds

$$\delta(\mathbf{x}) = \frac{\text{dist}\left((0,0,1), \mathbf{F}_{\text{ref},\hat{T}_i}^{-1}(\mathbf{x})\right)}{\text{dist}\left((0,0,1), \mathbf{Z}_{\text{ref}}^2(\mathbf{F}_{\text{ref},\hat{T}_i}^{-1}(\mathbf{x}))\right)}. \tag{7.5}$$

The next step in the construction of $\hat{\mathbf{\Phi}}_h$ is the same as in Section 2.3: connect $\mathbf{Z}_i^2(\mathbf{x})$ to a point on Γ_i and therefore, map the point $\mathbf{Z}_i(\mathbf{x})$ back onto the domain of the chart with $\hat{\boldsymbol{\gamma}}$. As $\hat{\boldsymbol{\gamma}}$ is invertible, define $\hat{\boldsymbol{\zeta}} = \boldsymbol{\gamma} \circ \hat{\boldsymbol{\gamma}}^{-1}$, such that $\hat{\boldsymbol{\zeta}}$ maps $\hat{\Gamma}$ on Γ. Then, the point $\hat{\boldsymbol{\zeta}}(\mathbf{Z}^2(\mathbf{x}))$ has to be mapped back onto the interior of the tetrahedron. As the mapping $\hat{\mathbf{\Phi}}_{h,i} = \hat{\mathbf{\Phi}}_{h|_{\hat{T}_i}}$ has to be the identity map on the faces of $\hat{T}_i$ that are not mapped on the curved boundary, the point $\hat{\boldsymbol{\zeta}}(\mathbf{z}(\mathbf{x}))$

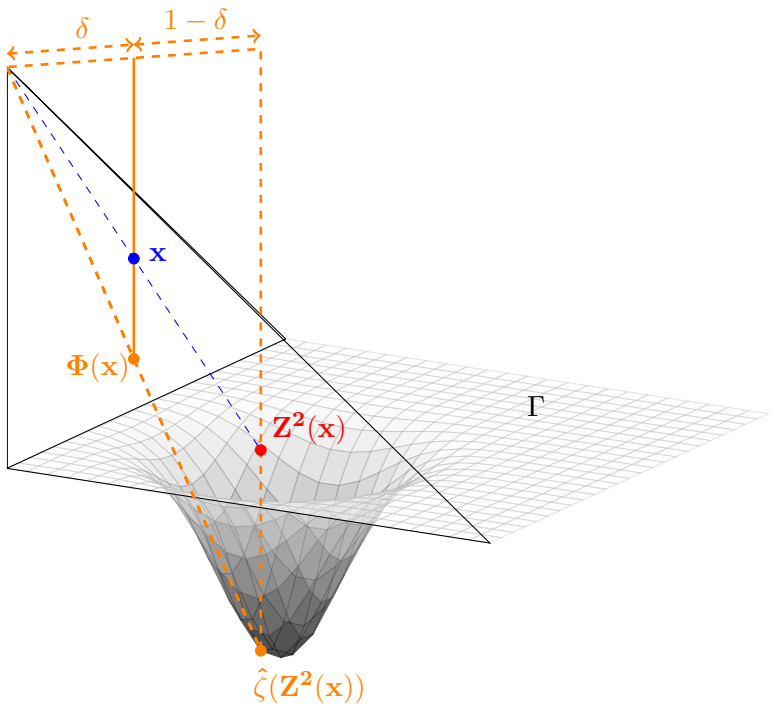


Figure 7.3: Construction of the mapping $\hat{\mathbf{\Phi}}_i,\ i \leq N$

has to be mapped back to the line throught the fourth point of the tetrahedron, and such that the ratio of the distance between $\mathscr{V}_3(\hat{T}_i)$ and $\hat{\boldsymbol{\zeta}}(\mathbf{Z}_i(\mathbf{x}))$ to the distance between $\mathscr{V}_3(\hat{T}_i)$ and the mapped back point is equal to $\delta(\mathbf{x})$. This is illustrated in Figures 7.2 and 7.3. Let $\mathbf{Y}^2_{i,\delta} : \Gamma_i \to \widetilde{T}_i$ denote this mapping:

$$\mathbf{Y}^2_{i,\delta} = \mathbf{F}^{-1}_{\text{ref},\hat{T}_i} \circ \mathbf{Y}^2_{\text{ref},\delta} \circ \mathbf{F}_{\text{ref},\hat{T}_i} \tag{7.6}$$

where

$$\begin{aligned} \mathbf{Y}^2_{\text{ref},\delta} &: \mathbb{R}^3 \to \mathbb{R}^3 \\ (x,y,z) &\mapsto (x\delta, y\delta, z\delta + 1 - \delta) \end{aligned} \tag{7.7}$$

Now, the mapping $\hat{\mathbf{\Phi}}_{h,i}$ can be defined as

$$\hat{\mathbf{\Phi}}_{h,i}(\mathbf{x}) = \mathbf{Y}^2_{i,\delta(\mathbf{x})}(\hat{\boldsymbol{\zeta}}(\mathbf{Z}^2_i(\mathbf{x}))),\ i \leq N. \tag{7.8}$$

For $N < i \leq \mathring{N}$ the point $\mathbf{x}$ on the interior of $\hat{T}_i$ has to be mapped onto E_i. This can be done considering the plane formed by $\mathbf{x}$ and $\bar{E}_i$ and choosing the unique intersection point with E_i, such that on each face of $\hat{T}_i$ whose intersection with Γ_i is E_i, this mapping is identical to $\mathbf{Z}^2$. This leads to

$$\begin{aligned} \mathbf{Z}^1_{\text{ref}} \ : \ \hat{T}_{\text{ref}} &\to \ E_{\text{ref}} \\ \mathbf{x}_{\text{ref}} &\mapsto \begin{cases} \left(-\frac{x_{\text{ref}}}{z_{\text{ref}}+y_{\text{ref}}-1}, 0, 0\right) & \mathbf{x} \notin \bar{E}_{\text{ref}} \\ \mathbf{x} & \mathbf{x} \in \bar{E}_{\text{ref}} \end{cases}. \end{aligned}$$

Similarly to the case $i \leq N$, define $\mathbf{Z}^1_i = \mathbf{F}_{\text{ref},\hat{T}_i} \circ \mathbf{Z}^1_{\text{ref}} \circ \mathbf{F}^{-1}_{\text{ref},\hat{T}_i}$. Then, the point $\hat{\boldsymbol{\zeta}}(\mathbf{Z}^1(\mathbf{x}))$ has to be mapped back onto the interior of the tetrahedron. Consider that the mapping $\hat{\mathbf{\Phi}}_{h,i}$ has to

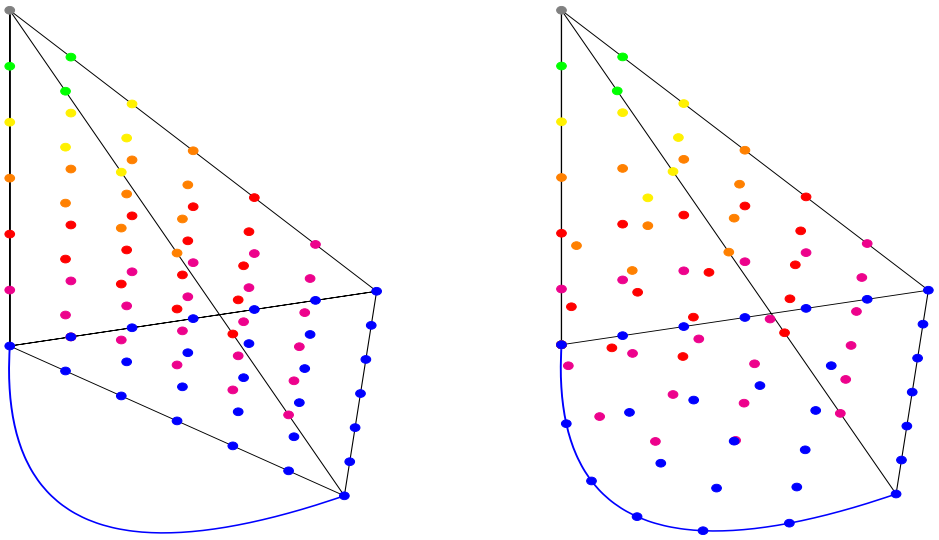


Figure 7.4: Example of the streching resulting from $\mathbf{Z}_i^1,\ N < i \leq \mathring{N}$

be the identity on the two faces of $\hat{T}_i$ whose intersection with $\hat{\Gamma}_i$ is not E_i and that it has to be identical with $\mathbf{Z}^2$ on the two other faces. This means that the barycentric coordinates of $\hat{\mathbf{\Phi}}(\mathbf{x})$ in the triangle formed with $\bar{E}_i$ and $\hat{\boldsymbol{\zeta}}(\mathbf{Z}^1(\mathbf{x}))$ have to be identical to the barycentric coordinates of $\mathbf{x}$ in the triangle formed with $\bar{E}_i$ and $\mathbf{Z}^1(\mathbf{x})$. Let (δ_1, δ_2) denote these barycentric coordinates. Then,

$$\begin{aligned} &\mathbf{Y}^1_{\mathrm{ref},\delta_1,\delta_2} : \mathbb{R}^3 \to \mathbb{R}^3 \\ &\qquad (x,y,z) \mapsto ((1-\delta_1-\delta_2)x, \delta_1 + y - \delta_1 y - \delta_2 y, \delta_2 + z - \delta_1 z - \delta_2 z)) \end{aligned} \tag{7.9}$$

maps a point $\mathbf{x}$ of $\mathbb{R}^3$ onto a point in the triangle formed by $\bar{E}_{\mathrm{ref}}$ and $\mathbf{x}$ at the barycentric coordinates (δ_1, δ_2). Note that the barycentric coordinates of $\mathbf{x}_{\mathrm{ref}}$ in the triangle formed with $\bar{E}_{\mathrm{ref}}$ and $\mathbf{Z}^1_{\mathrm{ref}}(\mathbf{x}_)$ are given by $(y_{\mathrm{ref}}, z_{\mathrm{ref}})$ such that

$$\hat{\mathbf{\Phi}}_{h,\mathrm{ref}}(\mathbf{x}) = \mathbf{Y}^1_{\mathrm{ref},y,z}(\hat{\boldsymbol{\zeta}}(\mathbf{Z}^1_{\mathrm{ref}}(\mathbf{x}))) \tag{7.10}$$

maps $\hat{T}_{\mathrm{ref}}$ onto $\widetilde{T}_{\mathrm{ref}}$. Using the transformation $\mathbf{F}_{\mathrm{ref},\hat{T}_i}$ leads to

$$\hat{\mathbf{\Phi}}_{h,i} = \mathbf{F}_{\mathrm{ref},\hat{T}_i} \circ \hat{\mathbf{\Phi}}_{h,\mathrm{ref}}(\mathbf{x}) \circ \mathbf{F}^{-1}_{\mathrm{ref},\hat{T}_i}\ . \tag{7.11}$$

This construction is illustrated in Figures 7.4 and 7.5. Setting $\hat{\mathbf{\Phi}}_{h,i} = \mathrm{id}$ for the interior elements leads to the global definition of $\hat{\mathbf{\Phi}}_h$. This construction leads to the same properties as in the two-dimensional case, i.e Theorem 2.1 holds (see [27]). For the construction of an approximated domain Ω_h with piecewise polynomial boundary, consider that the two-dimensional Lagrangian elements from Section 1.5 can be defined on the triangulation $\mathcal{S}_h$. Recall that the degrees of freedom are given by (see 1.44)

$$\mathcal{N} = \left\{ \boldsymbol{\gamma}(\mathfrak{F}_{i,l}) + \sum_{j=1}^{2} \frac{\lambda_j}{k} \left(\boldsymbol{\gamma}(\mathfrak{F}_{i,j}) - \boldsymbol{\gamma}(\mathfrak{F}_{i,l})\right) : 1 \leq l \leq 3,\ j \neq l, \lambda_j \in \mathbb{N}_0, \lambda_1 + \lambda_2 \leq k \right\} \tag{7.12}$$

for each $\mathfrak{F}_i$ in $\mathcal{S}_h$. Using the interpolation operator $\mathcal{I}_h$ with respect to these nodal points leads to parametrisation of the polynomial boundary $\gamma_h = \mathcal{I}_h \boldsymbol{\gamma}$. Let Ω_h denote the corresponding

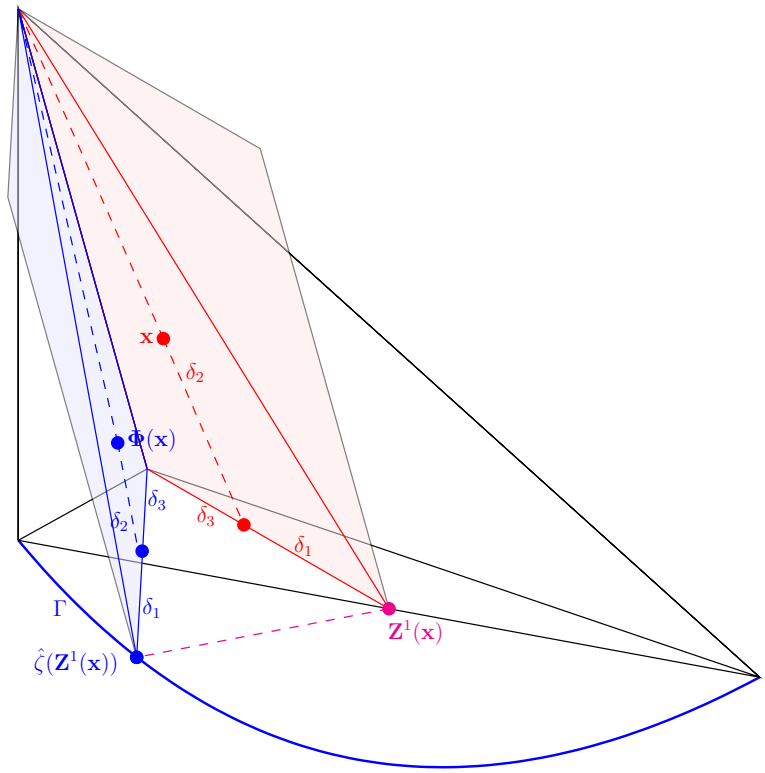


Figure 7.5: Construction of the mapping $\hat{\mathbf{\Phi}}_i,\ N < i \leq \mathring{N}$

polynomial domain and $\Gamma_h = \partial\Omega_h = \{\gamma_h(\mathbf{x}) \ : \mathbf{x} \in D\}$. $F_h : \hat{\Omega} \to \Omega_h$ can be constructed using the same way as above for $\hat{\mathbf{\Phi}}_h$, replacing $\boldsymbol{\gamma}$ by $\boldsymbol{\gamma}_h$, i.e. replacing $\hat{\boldsymbol{\zeta}}$ by $\boldsymbol{\zeta} = \boldsymbol{\gamma}_h \circ \hat{\gamma}_h^{-1}$

$$F_{h,i} = \mathbf{F}_{\text{ref},\hat{T}_i} \circ F_{h,\text{ref},i}(\mathbf{x}) \circ \mathbf{F}^{-1}_{\text{ref},\hat{T}_i} \tag{7.13}$$

with

$$F_{h,\text{ref},i} = \begin{cases} \mathbf{Y}^2_{\text{ref},\delta(\mathbf{x})}(\boldsymbol{\zeta}(\mathbf{Z}^2_{\text{ref}}(\mathbf{x}))) & i \leq N \\ \mathbf{Y}^1_{\text{ref},y,z}(\boldsymbol{\zeta}(\mathbf{Z}^1_{\text{ref}}(\mathbf{x}))) & N < i \leq \mathring{N} \\ id & i > \mathring{N} \end{cases} . \tag{7.14}$$

Similarly to the construction of $\hat{\mathbf{\Phi}}$, [27] proves that this construction leads to the same properties as in the two-dimensional case (Theorem 4.10). Moreover for the mapping $\mathbf{\Phi}_h = \hat{\mathbf{\Phi}}_h \circ F_h^{-1}$, Theorem 4.2 holds.

7.2 The Lagrange Element

Recall that the Lagrange element on an element T consists of scalar-valued polynomials of degree $k \geq 1$ and that the degrees of freedom are point values. As it is used to approximate $H^1(\Omega)$, the points have to be chosen in order to ensure interelement continuity. Therefore, $k+1$ points have to be placed on each edge of T and on each face F (see 1.43),

$$\dim \mathcal{P}_k(F) = \frac{1}{2}(k+1)(k+2), \tag{7.15}$$

points have to be placed. The standard choice is the usage of equidistant points such that on each face of T the degrees of freedom corresponds to the disposition in the two-dimensional case. Due to the fact that in the three-dimensional case

$$\dim \mathcal{P}_k(T) = \binom{k+3}{3} , \tag{7.16}$$

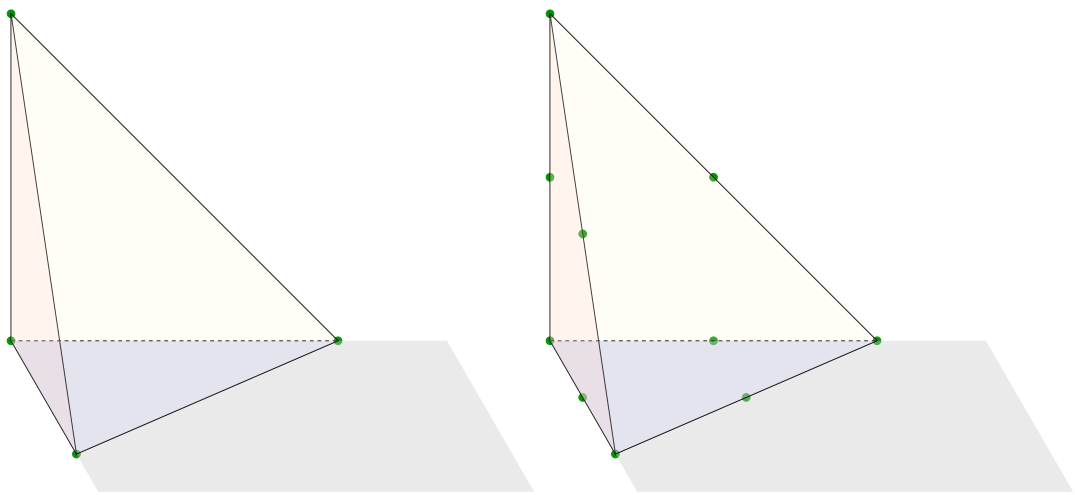
it remains

$$\begin{aligned} \mathcal{N}_I &= \frac{1}{6}(k+1)(k+2)(k+3) - 4\left(\frac{1}{2}(k-1)(k-2)\right) - 6\,(k-1) - 4 \\ &= \frac{1}{6}(k-1)(k-2)(k-3) \end{aligned} \tag{7.17}$$

nodes to be placed on the interior of the tetrahedron, for $k > 2$. For $k \leq 2$, there is no interior point, as illustrated in Figure 7.6. Note that

$$\mathcal{N}_I = \sum_{j=1}^{k-1} \frac{1}{2}(j-1)(j-2) \tag{7.18}$$

such that these degree of freedom can be placed similarly to the two-dimensional case, i.e. considering for each face of T the $k-1$ parallel triangles formed by the nodes on the edges that does not belong to this face.



(a) $k = 1$ (b) $k = 2$

Figure 7.6: Degrees of freedom for the Lagrange element ($d = 3$)

The interpolation operator is defined in the same way as in the two-dimensional case and the bound for the interpolation error (1.47) holds for $p \geq 2$.

7.3 The Raviart-Thomas Element

The space $RT_k(T)$ of the vector-valued Raviart-Thomas functions on a tetrahedron T is a subspace of $(\mathcal{P}_{k+1}(T))^3$ and the Raviart-Thomas finite-element space $RT_k(\Omega, \mathcal{T}_h)$ is a subspace of

the space of piecewise polynomials respect to T_h in each dimension:

$$RT_k(\Omega, \mathcal{T}_h) \subset \{\mathbf{v}_h \in (L^2(\Omega))^3 : \mathbf{v}_h|_T \in RT_k(T)\ \forall T \in \mathcal{T}_h\} \tag{7.19}$$

The sufficient condition for $H(\text{div}, \Omega)$-conformity in the three-dimensional case is that the normal component is continuous across each face that belongs to two elements, see [11]. The definition of the Raviart-Thomas element

$$RT_k(T) = (\mathcal{P}_k(T))^3 + \mathbf{x}\mathcal{P}_k(T) \tag{7.20}$$

ensures that enough degrees of freedom can be defined as the fluxes across faces of the mesh, such that these $H(\text{div}, \Omega)$-conformity holds. In fact, for $\mathbf{v}_h \in RT_k(T)$, it holds

$$\begin{aligned} \text{div } \mathbf{v}_h &\in \mathcal{P}_k(T) \\ \mathbf{v}_h \cdot \mathbf{n} &\in R_k(\partial T) \end{aligned} \tag{7.21}$$

with

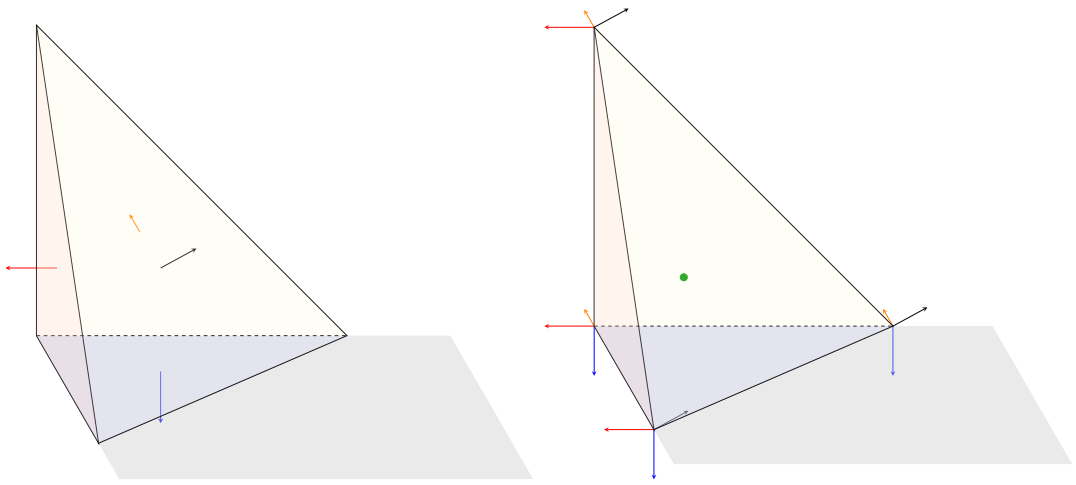
$$R_k(\partial T) = \{\phi \in L^2(\partial T) : \phi|_f \in \mathcal{P}_k(f) \quad \forall\ f \text{ face of } T\}\ . \tag{7.22}$$

The degrees of freedom $\mathcal{N}_{RT,T}$ are given by

$$\int_{\partial T} (\mathbf{n} \cdot \mathbf{v}_h) p_k\ ds, \quad p_k \in R_k(\partial T) \tag{7.23a}$$

$$\int_T \mathbf{v}_h \cdot \mathbf{p}_{k-1}\ d\mathbf{x}, \quad \mathbf{p}_{k-1} \in (P_{k-1}(T))^3 \tag{7.23b}$$

similarly to the two-dimensional case (see (1.56)). This is illustrated in Figure 7.7 for $k = 0$ and $k = 1$.



(a) $k = 0$ (b) $k = 1$

Figure 7.7: Degrees of freedom for the Raviart-Thomas element ($d = 3$)

For $k \geq 1$, the points for (7.23b) can be chosen as the interior points of the Lagrange element of type $k + 2$, each point represent 3 degrees of freedom, such that

$$\dim\ RT_k(K) = \frac{1}{2}(k+1)(k+2)(k+4)\ . \tag{7.24}$$

The definition and the properties (1.59) for the Raviart-Thomas interpolation operator holds in the three-dimensional case, see [11].

7.4 Numerical Results

This section presents the numerical experiments with a three-dimensional version of the example described in Section 3.3. The convergence results are now illustrated on the unit ball $\Omega = \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 < 1\}$. The formulation of the corresponding exact solution is identical to (3.42), only the constant α differs (see Figures 7.8 and 7.9)

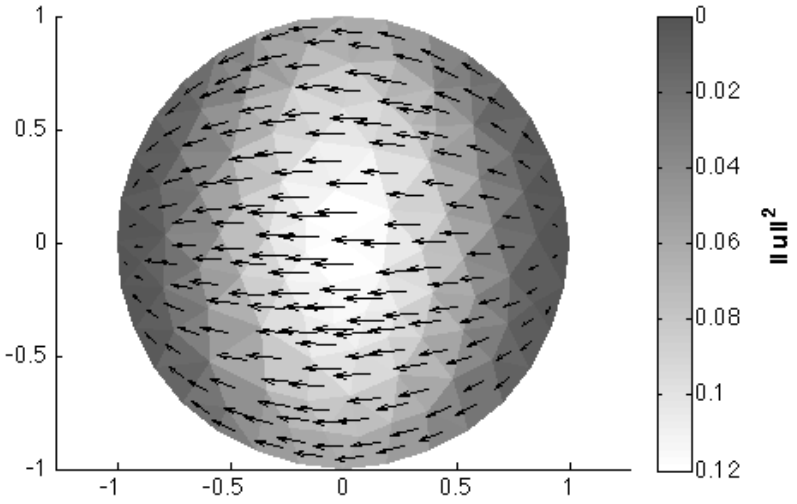


Figure 7.8: Approximate solution $\mathbf{u}_h$ on coarsest mesh

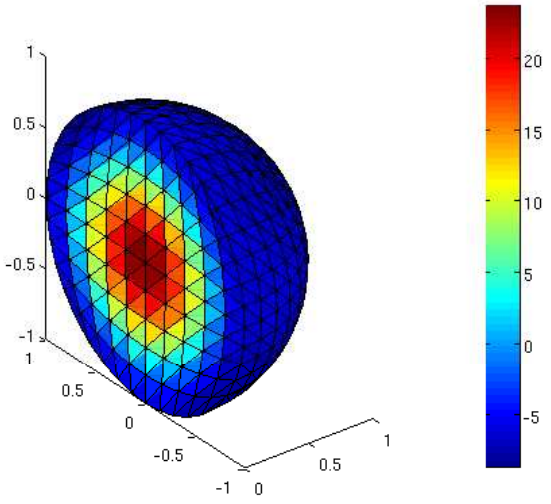


Figure 7.9: Exact solution p and triangulation on the unit ball

The results in Figure 7.10 show that the functional behaves in an optimal way proportionally to h^2 in the lowest-order case, as in the three-dimensional case it holds

$$h^2 \sim N^{-2/3} \,. \tag{7.25}$$

The square norm of the error is presented in Figure 7.11 and behave proportionally to h^2 in the lowest-order case as well. The convergence rate in the higher-order case (RT_1 combined with P_2 elements) is faster but the optimal convergence order 2 is not achieved. This is due to the inaccurate representation of the boundary, and similarly to the two-dimensional case a better approximation of the boundary and parametric elements on the domain with the interpolated boundary are needed to use higer-order elements. Their construction will be the topic of further work.

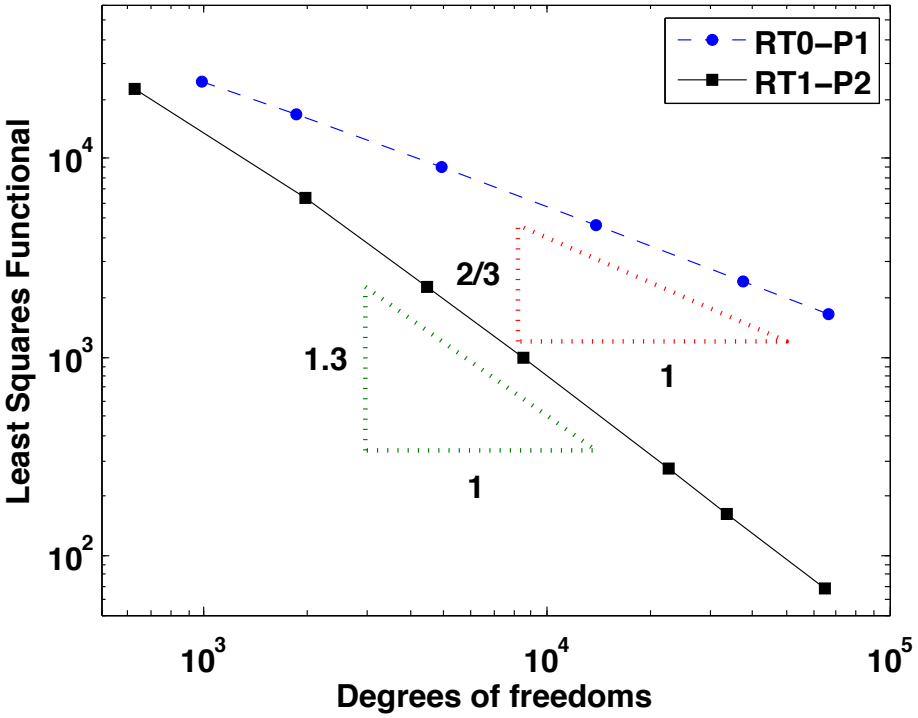


Figure 7.10: Convergence rates for the functional in $\Omega \subset \mathbb{R}^3$

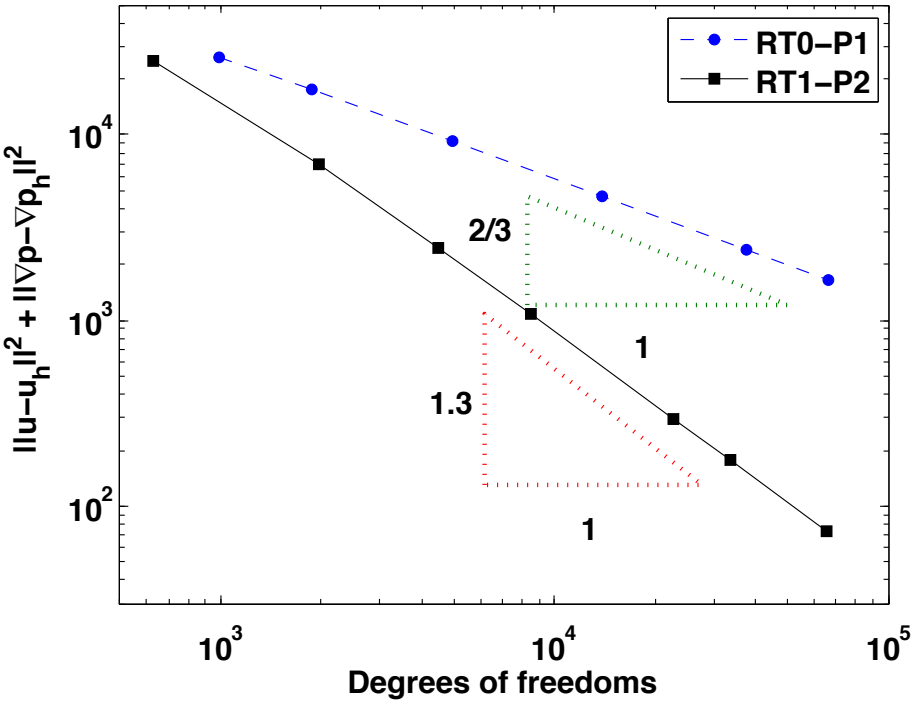


Figure 7.11: Convergence rates for the square norm of the error $\Omega \subset \mathbb{R}^3$

Conclusion and Outlook

In this thesis, optimal order convergence of a first-order system least squares method using (parametric) Raviart-Thomas elements combined with (isoparametric) conforming elements is shown for domains with curved boundaries. In particular, an estimate for the normal flux of (parametric) Raviart-Thomas elements on interpolated boundaries is derived. The computational results confirm the theory and an application to a two-phase model is given.

This framework can be used for the treatment of other problems involving a normal flux condition on a curved boundary or interface. This might include a lot of problems describing the motion of fluid substances, in particular two-phase flow. For example a least squares formulation for coupled Stokes-Darcy flow is derived in [33]. Since the typical least squares approach assumes that the nonlinear problem is a perturbation, this work suggests that nonlinear problems on curved boundaries could be analysed in the same way, applying the theory to the linearized equation. Moreover, two-phase problems with a time-dependent interface might be considered, i.e. free boundary problems, where the interface can be parametrized with a level set function (see [25] for a mixed method approach and [2] for a least squares approach). The coupled problems arising in water-mud interaction is an interesting application (see [22]).

Note that for problems with less regularity, triangulations which are not quasi-uniform can be used. The error estimator from the least squares approach allows an adaptive refinement which achieves better convergence rates in the presence of singularities. In this context an estimate of the form

$$|\langle \mathbf{n}\cdot\mathbf{v}_h, q\rangle_{0,\Gamma}| \lesssim \left(\|\mathbf{v}_h\|_{0,\Omega}^2 + \|\operatorname{div}\mathbf{v}_h\|_{0,\Omega}^2\right)^{1/2} \left(\sum_{i=1}^{N} h_i \|q\|_{1/2,\Gamma_T}^2\right)^{1/2}$$

would be used instead of (2.27).

A further point of interest is the approximation of the curvature due to the fact that some interface conditions depend on the curvature of the interface. The implementation of the three-dimensional parametric elements might be of interest for later research as well. Additional research is necessary to see if the estimates for the normal flux of the Raviart-Thomas elements on the interpolated boundaries derived in this work are useful in mixed finite-element methods or Hybrid-FOSLS as introduced in [29].

Bibliography

[1] R. A. Adams and J. F. Fournier, *Sobolev Spaces*, Academic Press, New York, 2nd ed., 2003.

[2] J. H. Adler, J. Brannick, C. Liu, T. Manteuffel, and L. Zikatanov, *First-order system least squares and the energetic variational approach for two-phase flow*, Journal of Computational Physics., v.230 n.17, p.6647-6663, July 2011.

[3] D. N. Arnold, J. Douglas and C. P. Gupta, *A family of higher order mixed finite element methods for plane elasticity*, Numerische Mathematik 1984.

[4] C. Bahriawati and C. Carstensen, *Three MATLAB implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control*, Computational Methods in Applied Mathematics, 5.4: 333–361, 2005.

[5] S. Bartels, C. Carstensen, and A. Hecht, *P2Q2Iso2D= 2D isoparametric FEM in Matlab*, Journal of Computational and Applied Mathematics 192.2: 219-250, 2006.

[6] S. Bauer, P. Neff, D. Pauly and G. Starke, *Dev-Div-and DevSym-DevCurl-inequalities for incompatible square tensor fields with mixed boundary conditions*, arXiv preprint arXiv:1307.1434, 2013.
or *Quelques inégalités de type Poincaré pour les champs de matrices quadratiques*, Comptes Rendus de l´Académie des Sciences - Mathématique 353(2):163-166, 2014.

[7] F. Bertrand, S. Münzenmaier, and G. Starke, *First-order system least squares on curved boundaries: Lowest-order Raviart-Thomas elements*, SIAM Journal on Numerical Analysis 52(2):880-894, 2014.

[8] F. Bertrand, S. Münzenmaier, and G. Starke, *First-order system least squares on curved boundaries: Higher-order Raviart-Thomas elements*, SIAM Journal on Numerical Analysis, (2013). Submitted for Publication.

[9] P. Bochev and M. Gunzburger, *Least-Squares Finite Element Methods*, Springer, New York, 2009.

[10] P. Bochev and M. Gunzburger, *On least-squares finite element methods for the Poisson equation and their connection to the Dirichlet and Kelvin principles.* SIAM Journal on Numerical Analysis 43.1: 340-362, 2005.

[11] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer, Heidelberg, 2013.

[12] D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, 2nd ed., 2001.

[13] J. Brandts, Y. Chen, and J. Yang, *A note on least-squares mixed finite elements in relation to standard and mixed finite elements*, IMA Journal of Numerical Analysis, 26:779–789, 2006.

[14] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 3rd ed., 2008.

[15] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part I*, SIAM Journal on Numerical Analysis, 31:1785-1799, 1994.

[16] Z. Cai, R. Lazarov, T. A. Manteuffel, and S. F. McCormick, *First-order system least squares for second-order partial differential equations: Part II*, SIAM Journal on Numerical Analysis, 34.2:425–454, 1997.

[17] C. Carstensen and D. Dolzmann, *A posteriori error estimates for mixed FEM in elasticity*, Numerische Mathematik, 81(2):187-209, 1998.

[18] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

[19] P. G. Ciarlet, *Mathematical Elasticity Volume I: Three–Dimensional Elasticity*, North-Holland, Amsterdam, 1988.

[20] M. Discacciati and A. Quarteroni, *Navier-Stokes/Darcy coupling: modeling, analysis and numerical approximation*. Revista Matematica Complutense, 22(2):315–426, 2009.

[21] A. Ern and J-L Guermond, *Theory and Practice of Finite Elements*, Springer, New York, 2004.

[22] J. Escher and A. V. Matioc, *On the well-posedness of a mathematical model describing water-mud interaction*, Mathematical Methods in the Applied Sciences, DOI: 10.1002/mma.2692 (online first)

[23] S. Ganesan, G. Matthies and L. Tobiska, *On spurious velocities in incompressible flow problems with interfaces*, Computer Methods in Applied Mechanics and Engineering 196(7): 1193-1202, 2007.

[24] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes Equations*, Springer, New York, 1986.

[25] S. Gross, V. Reichelt and A. Reusken, *A finite element based level set method for Two-Phase incompressible flows*, Computing and Visualization in Science, 9(4):239-257, 2006.

[26] S. Gross and A. Reusken, *Numerical methods for two-phase incompressible flows*, Springer-Verlag Berlin Heidelberg, 2011.

[27] M. Lenoir, *Optimal isoparametric finite elements and error estimates for domains involving curved boundaries*, SIAM Journal on Numerical Analysis, 23:562–58, 1986.

[28] J. Li, J. M. Melenk, B. Wohlmuth, and J. Zou, *Optimal a priori estimates for higher order finite elements for elliptic interface problems*, Applied Numerical Mathematics, 60:19–37, 2010.

[29] K. Liu, T. A. Manteuffel, S. F. Mccormick, J. W. Ruge, and L. Tang, *Hybrid first-order system least squares (hybrid-fosls) finite element methods with application to stokes equations*, SIAM Journal on Numerical Analysis, 51(4): 2214-2237, 2013.

[30] P. J. Matuszyk and L. F. Demkowicz, *Parametric finite elements, exact sequences and perfectly matched layers*, Computational Mechanics, 51:35–45, 2013.

[31] P. Monk, *Finite Element Methods for Maxwell's Equations*, Oxford University Press, New York, 2003.

[32] S. Münzenmaier, *First-Order System Least Squares for Generalized-Newtonian Coupled Stokes-Darcy Flow*, SIAM Journal on Numerical Analysis, (2014).
Submitted for Publication.

[33] S. Münzenmaier and G. Starke, *First-order system least squares for coupled Stokes-Darcy flow*, SIAM Journal on Numerical Analysis, 49:387–404, 2011.

[34] A. Pehlivanov, G. Carey and R. Lazarov, *Least-squares mixed finite elements for second-order elliptic problems*, SIAM Journal on Numerical Analysis 31.5: 1368-1377, 1994

[35] P.-A. Raviart and J.M. Thomas, *A mixed finite element method for 2-nd order elliptic problems*, Mathematical Aspects of Finite Element Methods, Springer Berlin Heidelberg, 1977.

[36] M. E. Rognes, R. C. Kirby, and A. Logg, *Efficient assembly of H(div) and H(curl) con- forming finite elements*, SIAM Journal on Scientific Computing, 31:4130–4151, 2009.

[37] R. Scott, *Interpolated boundary conditions in the finite element method*, SIAM Journal on Numerical Analysis 12.3:404–427, 1975.

[38] G. Starke, *Multilevel boundary functionals for least-squares mixed finite element methods*, SIAM Journal on Numerical Analysis 36:1065–1077, 1999.

[39] R. Stevenson, *A note on first order system least squares with inhomogeneous boundary conditions*, IMA Journal of Numerical Analysis, (2013). To Appear.

[40] J.-M. Thomas, *Méthode des éléments finis hybrides duaux pour les problèmes elliptiques du second ordre*, Revue française d´automatique, informatique, recherche opérationelle. Analyse numérique, 10(3):51–79, 1976

Name:	Fleurianne Bertrand
Adresse:	Nenndorfer Platz 70
	30459 Hannover
Email:	fleurianne.bertrand@gmx.de
Geburtsdatum:	6. Juni 1987
Geburtsort:	Caen, Frankreich
Nationalität:	französisch
Familienstand:	verheiratet

Studium und Ausbildung

Seit 11/2010 **Gottfried Wilhelm Leibniz Universität Hannover**
Wissenschaftliche Mitarbeiterin und Promotionsstudentin
- ▷ DFG-Projekt: Strömungsprozesse im Flüssigschlick: Mathematische Modellierung und numerische Simulation

10/2004-09/2010 **Gottfried Wilhelm Leibniz Universität Hannover**
Studium der Mathematik mit Studienrichtung *Rechnergestützte Wissenschaften* und dem Nebenfach Meteorologie:
- ▷ Gegenstand der Diplomarbeit: Semi-Lagrange-Methode mit transparenten Randbedinungen für Flachwassergleichungen
- ▷ Note: 1.3 (sehr gut)

09/2001-06/2004 **Lycée Blanche de Castille (Nantes, Frankreich)**
Abschluss: Baccalaureat, Schwerpunktfach Mathematik
- ▷ Endnote: 15.8/20

Veröffentlichungen

01/2014 SIAM Journal on Numerical Analysis 52(2) pp. 880-894, 2014:
First-Order System Least Squares on Curved Boundaries: Lowest-Order Raviart-Thomas Elements , mit Gerhard Starke und Steffen Münzenmaier

04/2013 Eingereicht: *First-Order System Least Squares on Curved Boundaries: Higher-Order Raviart-Thomas Elements*, mit Gerhard Starke und Steffen Münzenmaier

Betreute Lehrveranstaltungen

WS 2010/11 Übungen zu *Numerik partieller Differentialgleichungen*
SS 2012 Übungen zu *Gemischte partielle Differentialgleichungen*
WS 2012/13 Übungen zu *Numerik partieller Differentialgleichungen*
WS 2012/13 Seminar *Numerik nichtlinearer Probleme*

Ausseruniversitäre Erfahrungen

seit 08/2009 **ConsulPlex UG**
- ▷ Geschäftsführung und administrative Aufgaben
- ▷ Entwicklung von VBA Programmen
- ▷ Durchführung von VBA Workshops

07/2008-07/2009 **Allolio & Konrad Partnerschaft**
Praktika in Gatwick (UK), Paris (FR) und Bonn (GER)
- ▷ Abgleich von buchhalterischen Daten mit einem technischen Bestandsführungssystem
- ▷ Betreiben und Warten der Datenbanken (Microsoft SQL Server 2005)
- ▷ Durchführung von Workshops mit Nutzern und Betreuern der Datenbanken

07/2007-09/2007 **BHW Hameln**
Praktikantin in der Betriebsabteilung
- ▷ Entwicklung von VBA Anwendungen
- ▷ Erstellung eines Internetauftritts

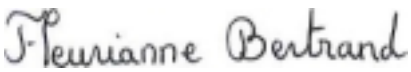
2006-2009 **Beratungsprojekte als studentische Unternehmensberaterin, u.a.:**
- ▷ Konzipierung und Erstellung einer Marktanalyse für eine Werbeagentur
- ▷ Prozessanalyse eines mittelständischen IT-Unternehmens
- ▷ Erstellung eines Handbuchs der unternehmenseigenen ERP-Software

Sprachkenntnisse

Französisch	Muttersprache
Deutsch	Fließend in Wort und Schrift
Englisch	Fortgeschrittene Kenntnisse
Latein	Großes Latinum

Hobbys & weitere Interessen

Sport	Skifahren
Sonstiges	Tenorflöte im Ensemble, VBA-Programmierung



Fleurianne Bertrand,
10.05.2014

Nenndorfer Platz 70, 30459 Hannover · E-Mail: fleurianne.bertrand@gmx.de